

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 16:48:39 ; Search time 4873.04 Seconds
(without alignments)
1604.679 Million cell updates/sec

Title: US-09-236-995D-3
Perfect score: 474
Sequence: 1 aacaagatgctattatggca.....gtttccatcacaaagaggtag 474

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_scs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_scs.*
- 27: em_sy.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgtgo_hum.*
- 31: em_hgtgo_inv.*
- 32: em_hgtgo_rod.*
- 33: em_hgtgo_hum.*
- 34: em_hgtg_inv.*
- 35: em_hgtg_rod.*
- 36: em_hgtg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
1	474	100.0	3211	8	ZMPARP2	AJ222589 Zea mays
2	474	100.0	3285	8	AF093627	AF093627 Zea mays
3	280	59.1	3187	8	ATH131705	AJ131705 Arabidops
4	158.6	33.5	1939	10	RNPARP2	X65497 R.norvegicu
5	157.6	33.2	3718	5	XELFADPRP	D14667 Frog mRNA f
6	155	33.1	3128	10	RNU94340	U94340 Rattus norv
7	157	32.7	3617	5	XLNAPRG	212139 X.laeviss PA
8	154.4	32.6	2147	8	ATPARP	248243 A.thaliana
9	153.8	32.4	3036	5	GGPADPRP	X52690 Chicken mRN
10	152.2	32.1	3845	10	BC012041	BC012041 Mus muscu
11	151.4	31.9	2295	8	ZMPARP1	AJ222588 Zea mays
12	150.6	31.8	3172	10	MMADPRP	X14206 Mouse mRNA
13	150.6	31.8	3589	4	BOVPAS	D90073 Bovine poly
14	149	31.4	1732	10	AF126717	AF126717 Mus muscu
15	147.4	31.1	3047	10	AF168781	AF168781 Crictellu
16	142.6	30.1	1771	9	HUMFOLP	M17081 Human place
17	142.6	30.1	3640	9	HUMFOLP	M18112 Human poly(
18	142.6	30.1	3792	6	A52134	A52134 Sequence 1
19	142.6	30.1	3795	9	HUMRISDAD	J03473 Human poly(
20	142.6	30.1	5787	9	GORADPRB	I24094 Gorilla gor
21	139.4	29.4	3045	6	AX058340	AX058340 Sequence
22	139.4	29.4	3045	6	AX062277	AX062277 Sequence
23	139.4	29.4	3660	9	HUMFOLP	M32721 Human poly(
24	139.4	29.4	3747	6	I14359	I14359 Sequence 5
25	137.8	29.1	1822	3	DMPARP5	AF051548 Drosophil
26	137.8	29.1	2682	9	S54638	S54638 (pseudogene
27	137.8	29.1	3080	3	DROADPRPA	D13806 Fruit fly m
28	137.8	29.1	5345	6	I14360	I14360 Sequence 7
29	137.8	29.1	9254	9	HUMFADPRP2	L14752 Homo sapien
30	137.8	29.1	10988	2	AC014613	AC014613 Drosophil
31	137.8	29.1	16864	3	AE002892	AE002892 Drosophil
32	137.8	29.1	155026	9	AL442128	AL442128 Human DNA
33	137.4	29.0	3463	3	SPEPADPRP	D16482 Sarcophaga
34	136.2	28.7	2682	6	I14357	I14357 Sequence 3
35	122.6	25.9	687	6	AX058382	AX058382 Sequence
36	122.6	25.9	738	6	AX058380	AX058380 Sequence
37	122.6	25.9	864	6	AX058336	AX058336 Sequence
38	122.6	25.9	1754	9	HSA236912	AJ236912 Homo sapi
39	122.6	25.9	1797	9	HSA236876	AJ236876 Homo sapi
40	122.6	25.9	1814	6	AX058317	AX058317 Sequence
41	122.6	25.9	1843	6	AX008892	AX008892 Sequence
42	122.6	25.9	1874	6	AX058374	AX058374 Sequence
43	122.6	25.9	3200	6	AX058362	AX058362 Sequence
44	121	25.5	1910	9	AF085734	AF085734 Homo sapi
45	120.2	25.4	1707	6	AX058338	AX058338 Sequence

ALIGNMENTS

RESULT	1	ZMPARP2	ZMPARP2	3211 bp	mRNA	PLN	19-NOV-1997
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							

JOURNAL Submitted (14-NOV-1997) Kushnir S., VIB, Dep.Genetics,
Ledeganckstraat 35, Gent, B9000, Belgium

FEATURES
Source 1...3211
Location/Qualifiers
/organism="Zea mays"
/db_xref="taxon:4577"
113..3022
/gene="PARP"
113..3022
/gene="PARP"
/function="secondary protein modification"
/codon_start=1
/product="poly(ADP-ribose) polymerase"
/protein_id="CA10889.1"
/db_xref="GI:2632129"
/db_xref="SPTRMBL:O24570"
/translation="MAAPPKAWKAEYAKSGRASCSCRSPIAKDQLRLGKMWQASQFD
GFPMNHNHARSIFSKKNDKSVDDVEGIDALRWDDQEKIRNYVSASAGTSSTAAPPEKTIETAPSA
TSCRCSEKIKTGSRVLSAKLESEKPGIPWYHANGFFEVSPSATVEKFGMDTLSD
DKRTMLDLVKDNGNNEONKSGSKRSKSENDIDSKYKARLDESECTVKNKOLVDPR
GSNTSSADIGLPCPCANGMYIYNGQYCCSGNSVSEKCTISATEPRVKKKWOIPHGT
KNDYLMKFKSKQKVPKPERVLPMSPEKSGSKATQRTSLSSKGLDKLRFVVGOSKE
AANEIEKLKAGANFYRVKIDICLIAGELDNENAERYKARLKIPIVREGYIGE
CVKNKMLPDLKYLENALESSKSTVTVKVRGSAVHESGLQDTAHILEDKSYIYN
ATLMSDLALGVNSYVLOIEEODGSECVFRKGRVSEKIGGOKLEMSKTEAIK
EFKRLFLKGTNSWEAWECKTNFRKOPGRFYPLDVOYGVKKAPKRDISEMSSLAPO
LLEMLKMFNVEYRAMMEFEINMSEPLGKLSKENIEKGFALTEIONLLKDTAQ
ALAVRESLIVAASNRFTLIPSIHPHIIREDDELMIKAKMLEALQDIEASKIVGDS
DSDESDDKYMKLHCDITPLAHDSYKLIQYLLNTHAPTHKDSLEEVFSLDRD
GELNYSRYKNLHNMKLLHWSRLTNFVGLISQGLRIAPPAPVTGYMFGKGLYFAD
LVSKSAQCYVDRNPNVGLMLSEVALGDMYELKATSMDKPPRGKHSTKGLKTVPL
ESEFVKWRDDVVVPCGKPVPSIRSELNMYEIVNTSQVKMQFLLLKVRFHKR"

BASE COUNT 968 a 604 c 813 g 826 t

ORIGIN

Query Match 100.0%; Score 474; DB 8; Length 3211;
Best Local Similarity 100.0%; Pred. No. 9.2e-139; Indels 0; Gaps 0;
Matches 474; Conservative 0; Mismatches 0;

Qy 1 acaagatgctattatggcaggttcgaagttgacgaatttgggaattcttagtcaa 60
Db 2549 AACAGATGCTATTATGCGACGGTTCAAGGTTGACGAATTTTGTGGAAATCTTAGTCAA 2608

Qy 61 gggctaagaattgcaactctctgagcaccctgttactggttatgttcggcaaggcctc 120
Db 2609 GGGCTAAGAATTGCACTCTCTGAGGCACCTGTTACTGGCTATATGTTTCGGCAAGGCCTC 2668

Qy 121 tactttgcagatctagtaagaagagcgacacatactgttatgtgataggaataatcct 180
Db 2669 TACTTTGCAGATCTAGTAAGAAGAGCGGCAACATACHTGTTATGTGATAGGAATAATCCT 2728

Qy 181 gtaggtttgatgctcttcttctgaggttcttaggagacatgtatgaactaaagaagcc 240
Db 2729 GTAGGTTTGATGCTCTCTTCTGAGGTGCTTTAGGAGACATGTATGAACATAAGAGCC 2788

Qy 241 acgtccatgagcaaacctccaaagaggaagcattcagcaccagggataggcaaaacccgtg 300
Db 2789 ACGTCCATGAGCAAACTCCAAAGAGGGAAGCATTCGACCAAGGATTAGGCAAAACCGGTG 2848

Qy 301 ccactgagtcagagttgtgagtgagggatgagtgatgctgtagtccctcgagcaagccg 360
Db 2849 CCACTGAGTCAGAGTTTGTGAAGTGAGGGATGATGCTGATGTTCCCTCGGCAAGCCG 2908

Qy 361 gtgccatcatcaattagagagctctgaactcatgtacaatgagtcacatcgctcacacaca 420
Db 2909 GTGCCATCATCAATTAGGAGCTCTGAACCTCATGTACAATGAGTACATCGTCTACACACA 2968

Qy 421 tccagggtgaagatgcaagttcttctgtaaggtgctgttccatcacagaaggtgag 474
Db 2969 TCCAGGTGAAGATGAGTCTTCTGCTGAAGGTGCGTTTCCATCACAGAAGGTAG 3022

RESULT 2

AF093627 3285 bp mRNA PLN 29-NOV-1998
Zea mays poly(ADP)-ribose polymerase (PARP1) mRNA, complete cds.
AF093627
AF093627.1 GI:3928870
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3285)
Mahajan, P.B. and Zuo, Z.
Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase
Plant Physiol. 118 (3), 895-905 (1998)
99026291
2 (bases 1 to 3285)
Mahajan, P.B. and Zuo, Z.
Direct Submission
Submitted (22-SEP-1998) Crop Protection, Pioneer HiBred
International, Inc., 7300 NW 62nd Avenue, Johnston, IA 50131-1004,
USA
Location/Qualifiers
1..3285
/organism="Zea mays"
/db_xref="taxon:4577"
1..3285
/gene="PARP1"
100..3042
/gene="PARP1"
/EC_number="2.4.2.30"
/function="add ADP-ribose"
/codon_start=1
/product="poly(ADP)-ribose polymerase"
/protein_id="AAC75704.1"
/db_xref="GI:3928871"
/translation="MAAPPKAWKAEYAKSGRASCSCRSPIAKDQLRLGKMWQASQFD
GFPMNHNHARSIFSKKNDKSVDDVEGIDALRWDDQEKIRNYVSASAGTSSTAAPPE
KTIETAPSAATSCRCSEKIKTGSRVLSAKLESEKPGIPWYHANGFFEVSPSATVE
KFGSDTLSDDEKRTMLDLVKDNGNNEONKSGSKRSKSENDIDSKYKARLDESECT
VKNKOLVDPRGSGNTSSADIGLPCPCANGMYIYNGQYCCSGNSVSEKCTISATEPR
VKKMQIPHGTNDYLMKFKSKQKVPKPERVLPMSPEKSGSKATQRTSLSSKGLDK
LRFVVGOSKEAANEIEKLKAGANFYRVKIDICLIAGELDNENAERYKARLKIPIV
REGYIGEVCVKNKMLPDLKYLENALESSKSTVTVKVRGSAVHESGLQDTAHILE
DKSYIYNATLMSDLALGVNSYVLOIEEODGSECVFRKGRVSEKIGGOKLEMSKTE
AIKEFKRLFLKGTNSWEAWECKTNFRKOPGRFYPLDVOYGVKKAPKRDISEMSS
LAPOLEMLKMFNVEYRAMMEFEINMSEPLGKLSKENIEKGFALTEIONLLKDTAQ
ALAVRESLIVAASNRFTLIPSIHPHIIREDDELMIKAKMLEALQDIEASKIVGDS
EIEEVFSLDRDGELENNKYSRYKNLHNMKLLHWSRLTNFVGLISQGLRIAPPAPVT
GYMFGKGLYFADLVSKSAQCYVDRNPNVGLMLSEVALGDMYELKATSMDKPPRGKH
STKGLKTVPLESEFVKWRDDVVVPCGKPVPSIRSELNMYEIVNTSQVKMQFLLL
KVRFHKR"

BASE COUNT 1003 a 617 c 828 g 837 t

ORIGIN

Query Match 100.0%; Score 474; DB 8; Length 3285;
Best Local Similarity 100.0%; Pred. No. 9.3e-139; Indels 0; Gaps 0;
Matches 474; Conservative 0; Mismatches 0;

Qy 1 acaagatgctattatggcaggttcgaagttgacgaatttgggaattcttagtcaa 60
Db 2569 AACAGATGCTATTATGCGACGGTTCAAGGTTGACGAATTTTGTGGAAATCTTAGTCAA 2628

Qy 61 gggctaagaattgcaactctctgagcaccctgttactggttatgttcggcaaggcctc 120
Db 2629 GGGCTAAGAATTGCACTCTCTGAGGCACCTGTTACTGGCTATATGTTTCGGCAAGGCCTC 2688

Qy 121 tactttgcagatctagtaagaagagcgacacatactgttatgtgataggaataatcct 180

```
|||||
Db 2689 TACTTTGCAGATCTAGTAAGCAAGAGCGCACATACTGTATGTGATAGGAATAATCCT 2748
QY 181 gtaggttatgattcttctcaggttctttaggagacatatgaactaagaagacc 240
Db 2749 GTAGGTTTGTATGCTTCTTCTGAGGTTGCTTTAGGAGACATGTATCAACTAAGAAAGCC 2808
QY 241 acgtccatgacaaacctccaagaggaagcattcgaccagggattagcacaacccgtg 300
Db 2809 ACGTCCATGACAAACCTCCAGAGGGAAGCATTCGACCAAGGATTAGGCAAAACCGGTG 2868
QY 301 coactgagtcagagttgtgagtgagggatgtatgtctgagttccctgcggcaagccg 360
Db 2869 CCAGTGGAGTCAGAGTTTGTGAGTGGAGGATGTATGCTGATTTCCCTGCGGCAAGCGG 2928
QY 361 gtccatcatcaattaggagctcctgaactcatgtcaatgagtcacatgctcacaacaca 420
Db 2929 GTCCCATCATCAATAGGAGCTCTGAACATCATGTACATGATGATGATGATGATGATGAT 2988
QY 421 tccaggtgaagtcaggttcttctgtaaggtgcgtttccatcacaagaggtag 474
Db 2989 TCCAGGTGAAGTCAGTCTTCTGCTGAAGGTGCGTTTCCATCACAAGAGGTAG 3042

RESULT 3
ATH131705 3187 bp mRNA PLN 18-DEC-1998
LOCUS Arabidopsis thaliana mRNA for poly(ADP-ribose) polymerase.
DEFINITION AJ131705
ACCESSION AJ131705
VERSION AJ131705.1 GI:4038490
KEYWORDS NAD(+) ADP-ribosyltransferase; parp-1 gene; poly(ADP-ribose) polymerase.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 3187)
Direct Submission
Submitted (17-DEC-1998) Kazmaier M., Departement D'Ecophysiologie Vegetale Et De Microbiologie, Commissariat A L'Energie Atomique, CEA Cadarache, Lab. de Radiobiol. Vegetale, Bat. 185, 13108 St. Paul-les-Durance, FRANCE
Paul-les-Durance, FRANCE
Doucet-Chabeaud, G. and Kazmaier, M.
Unpublished
Location/Qualifiers
1..3187
/organism="Arabidopsis thaliana"
/cultivar="landsberg erecta"
/db_xref="taxon:3702"
11..2963
/gene="parp-1"
11..2963
/gene="parp-1"
/standard_name="NAD(+) ADP-ribosyltransferase"
/EC_number="2.4.2.30"
/function="ADP-Ribose polymer synthesis"
/notes="mol. weight: 111 kDa"
/codon_start=1
/evidence-experimental
/product="poly(ADP-ribose) polymerase"
/protein_id="CA010482.1"
/db_xref="GI:4038491"
/translation="MASPHKPRAEVAKSSRSCKTKSVINKENFRGLKLVOSTHFD
GIMPWNHASCILIKTKQIKSVDDVEGIESLRWEDQOKIRKYVESGASNTSTGTS
TSTANNHAGIEVYSQTRAGRCSEKILKVEVIRIPKPEGNGKLMWHAKCF
LEMSSSTELSLGWSRIPDSQEAALLPLVKALPAKTAETAEAROTYSRAGTKRND
SVNKSLEKLSFMSFGALQPCSKEMEAOQKELMDLKKYVTSALREML
EVNEQSTRGSELDLRDKCADGMFGPLALCPMSGHLSPGLYRCHGYISWKSCH
STLDPRIKWKIPDETENQFLKWNKSQKSVKPKRILRPVLSGTSOGGSKATD
SSKSEKLADLKLVISAGNTRKOPWRKRIEAGAEFANVKKGTSLVVCGLDIRDAE"

REFERENCE
AUTHORS Penning, C., Poirier, G.G. and Moreau, P.
TITLE Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase catalytic domain and analysis of mRNA levels during the cell cycle
JOURNAL Biochem. Cell Biol. 67 (9), 653-660 (1989)

MRKARMKVAIVREDYLVDCFKQKLPDKYKIEDTSLSLTVKVKRSVHSEASGL
GNEHLEDGNSYNTLSNDSLTGINSYIILOIEDKSGDCYVFRKWRGVNGEKI
GGKPEMSKSDAVHEFKRLMLEKNTGWESWQKTFQKPGKFLPIDGTVNKKV
AKPEPOTSSNLAPSLIELMLKLVDETVRSAMMEFEINMSEPLKLSKHNIQKGE
ALTEIORLLTESDPOTMKESSLVDASNRFTWIPSTPHIIRDEDDFKSKVLMLEAL
ODIEASRIYGVDFDSTESLDKKYKLLHCDISPLPHDSEYRLIEKLYLTHAPTHTE
WSLELEVFALEREGEREDKYAPREKLGKMLLHGSRLTNFVGLINQGLRTAPPEAP
ATGYMFGKIYFADLVSKSAOYCTCKNPNVGLMLLSEVALGEIHELTAKAKYMDKPPR
GKSTGLGKGVQDSEFAKRWGDTVPCKPVSSVKASELMYNEYIYVDTAQVKLQ
FLKVRFPKHKR"
14..2959
mat_peptide
/gene="parp-1"
/EC_number="2.4.2.30"
/product="poly(ADP-ribose) polymerase"
BASE COUNT 1044 a 577 c 743 g 823 t
ORIGIN

Query Match 59.18; Score 280; DB 8; Length 3187;
Best Local Similarity 74.6%; Pred. No. 2.1e-77;
Matches 352; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 1 aacaagatgcttatatggcagcgttcacgaatttgggaatttcttagtcaa 60
Db 2489 AATAAGATGCTCCTATGGCATGGTCTCGATTACGAATTTTGTGGATATTGAACAA 2548
QY 61 gggctaaagattgcacctcctgagcgacctgttactggctatgttctggcaggcctc 120
Db 2549 GGACTGAGAATTGCACTCCAGAAAGCTCTGCTACTGTTACATGTTTGGAAAAAGGATA 2608
QY 121 tactttgagatcttagtaagcaagagcgcacaaactactgttattgttgatagaataactc 180
Db 2609 TACTTTGCTGACCTTGTGAGTAAAGTCTCAGTACTGCTACACTTGTAAAGAAAAATCCG 2668
QY 181 gtaggttttagtctcttctgaggttctttaggagacatgtatgaactaaagaaagcc 240
Db 2669 GTGGGTCTAATGCTTCTCAGTGAAGTTCATTTGGAGAAATACATGAGTAAACAAAGCT 2728
QY 241 acgtccatgacaaacctccaagagggagcattcgaccagggattaggcaaacccgtg 300
Db 2729 AAGTATATGGATAAACCTCCGAGAGGGGAAACACTCGACAAAGGCTCGGCAAGAAAGTG 2788
QY 301 coactgagtcagagttgtgaagtgagggatgatcgtagttccctgcggcaagccg 360
Db 2789 COTCAAGATTCGAGTTTGGCAAGTGAGAGGTGATGACTGTTTCCCTGTGGAAACCT 2848
QY 361 gtgccatcatcaattaggagctctgaactcatgtacaaatgagtcacatgctcacaacaca 420
Db 2849 GTTTCATCAAGTCAAGCTTCTGAGCTTATGTACATGATGATATGCTCTACGATACA 2908
QY 421 tccaggtgaagtcagttcttctgtaaggtgcgtttccatcacaagaggt 472
Db 2909 GCCCAGGTGAAGTTGCTTCTTGTGAAAGTAAAGGTTTAAAGCACACAGAGAT 2960

RESULT 4
RNPARP2 1939 bp DNA ROD
LOCUS R.norvegicus mRNA for poly(ADP-ribose) polymerase (clone PRATC).
DEFINITION X65497
ACCESSION X65497
VERSION X65497.1 GI:56849
KEYWORDS NAD(+) ADP-ribosyltransferase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1939)
REFERENCE 1 (bases 1 to 1939)
AUTHORS Thibodeau, J., Gradwohl, G., Dumas, C., Clairoux-Moreau, S., Brunet, G., Penning, C., Poirier, G.G. and Moreau, P.
TITLE Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase catalytic domain and analysis of mRNA levels during the cell cycle
JOURNAL Biochem. Cell Biol. 67 (9), 653-660 (1989)
```

```
MEDLINE
REFERENCE 90027702
AUTHORS Potvin,F.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-1992) F. Potvin, Molecular Endocrinology, CHUL
Research Center, 2705, Boul Laurier, Ste-Foy, Quebec, G1V 4G2,
CANADA
COMMENT See also X65496.
FEATURES Location/Qualifiers
SOURCE 1..1939
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
/chromosome="1q41-42"
/tissue_type="prostate"
/clone_lib="lambda EMBL4"
/clone="PRATC"
32..1510
/EC_number="2.4.2.30"
/codon_start=1
/product="Poly(ADP-ribose) polymerase"
/protein_id="CA46478.1"
/db_xref="GI:56850"
/translation="MKLTGKGAADVDPDGLSHAHVLEKGGKVFSAITGLDIVKGT
NSYKQLLESDEKSEWTFPSRSGRVGTGSKNLEBOMPSEKDAVEHFMKLYEKTGN
AWHSKNTKYPKPHPLAIDYGODEAVKVLAKPGTKSKLPKQVELGMFLDVESM
KKALEYEIDLQKPLKLSRQIQAYISLSEVOQAVSQGSESOILDLNRFYTLI
PHFGMDRPLEDNTDQVAKVEMLDNLDIEVAYSLRGSDSDSKPDIDNYPEKLL
TDIKVDROSEEAIEVKYKNTHTATTHAYDLEVIDIEFIEREGESQRYKPFROLHN
RLLMHSRTTNPAGILSGLRIAPPEAPVTGFMFGKGIYFADMYKSNYCHTSQGD
PILGLLGEVALGNMYELHSHASIKLPKSHSVKGLGTAPDPSASITLDGVEVPLG
TGIPSGVNTCLLYNEYIYVDIAOVNKLKLLKFNKFTSLW"
BASE COUNT 524 a 468 c 521 g 426 t
ORIGIN

Query Match 33.5%; Score 158.6; DB 10; Length 1939;
Best Local Similarity 60.1%; Pred. No. 4.8e-39;
Matches 282; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

Qy 1 acaagatgctattatggcaggttcaaggttgacgaatttgggaattcttagtcaa 60
Db 1031 AACCGGAGACTGCTGTGGCAGCGGTCCAGGACCAACCACTTCGACGATCTGTCCACAG 1090

Qy 61 gggttaagaattgcacctcttgaggaacctgttactggtctatatgttgcgaagccctc 120
Db 1091 GGCTCGGATAGCCCCACCTGAAGCACCTGTGACAGGCTACATGTTGGGAAAGGAATC 1150

Qy 121 tactttgcagatctagtaagaagagcgcaataactgttattgtgtagaataatactct 180
Db 1151 TACTTTGCTGATGGTGTCCAAAGTGGCAACTACTGCCACACGTCCTCAGGGAGACCCG 1210

Qy 181 gtaggtttgatgctcttcttggaggttggtttaggagacatgtatgaactaaagaagcc 240
Db 1211 ATTGGCTTAATCTGTTGGGAGAAAGTTGCCCTTGGAAACATGTACGAACCTCAAGCATGCT 1270

Qy 241 acgtccatgacaaacctccaagagggaagcattcgaccacgaaggattagcaaaaccgtg 300
Db 1271 TCTCAGATCAGCAAGTTTACCAAGGGCAAGCAGACAGTGTCAAAGGTTTGGGCAAAACCG-- 1328

Qy 301 ccactggagtcagagtttggagtgaggagtgatgtctgttagttccctgcggaagccg 360
Db 1329 -CCCTGACCTTCGGCCACCATCACCTCGGATGGTGTAGAGGTTCCGCTGGGACAGCG 1387

Qy 361 gtgcatacataataggagctctgaactcatgtacacatgtagtacatctctacaaca 420
Db 1388 ATTCGGTCTGGTGTATGATGACACTGCTGCTGTATAAAGAGTACATTTGCTACGACATT 1447

Qy 421 tccaggtgaagatcagatcttgcgaaggtgcgtttccatcacaga 469
Db 1448 GCTCAGTGTAATCTGAAGTACCTACTGAAACTGAAGTTCAATTTCAAGA 1496
```

```
RESULT 5
XELPADPRP
LOCUS 3718 bp mRNA VRT 04-FEB-1999
DEFINITION Frog mRNA for poly(ADP-ribose) polymerase.
ACCESSION D14667
VERSION D14667.1 GI:287584
KEYWORDS poly (ADP-ribose) polymerase,
SOURCE Xenopus laevis egg, oocytes, germ cell, cDNA to mRNA, clone
XPAP5-9R.
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
Xenopodinae; Xenopus.
Uchida,K.
1 (bases 1 to 3718)
Direct Submission
Submitted (15-MAR-1993) to the DDBJ/EMBL/GenBank databases.
Kazuhiko Uchida, Institute of Basic Medical Sciences, Department of
Biochemistry, Ten-noudai 1-1-1, Tsukuba, Ibaraki 305, Japan
(E-mail:kzuchida@md.tsukuba.ac.jp, Tel:0298-53-3271,
Fax:0298-53-3039)
2 (bases 1 to 3718)
Uchida,K., Uchida,M., Hanai,S., Ozawa,Y., Ami,Y., Kushida,S. and
Miwa,M.
TITLE Isolation of the poly(ADP-ribose) polymerase-encoding cDNA from
Xenopus laevis: phylogenetic conservation of the functional domains
JOURNAL Gene 137 (2), 293-297 (1993)
MEDLINE 94131300
REFERENCE 3 (bases 1 to 3718)
AUTHORS Uchida,K., Uchida,M., Hanai,S., Ishikawa,K., Ozawa,Y., Ueno,N. and
Miwa,M.
TITLE Isolation of Poly(ADP-ribose) Polymerase cDNAs from Xenopus laevis
and Cherry Salmon using heterologous oligonucleotide consensus
sequences: gene conservation in amphibia, fish, and insect
Unpublished (1993)
Submitted (15-MAR-1993) to DDBJ by:
Kazuhiko Uchida
Department of Biochemistry
Institute of Basic Medical Sciences
University of Tsukuba
1-1-1 Ten-noudai
Tsukuba, Ibaraki 305
Japan
Phone: 0298-53-3271
Email: kzuchida@md.tsukuba.ac.jp
Fax: 0298-53-3039.
Location/Qualifiers
1..3718
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="Germ cell"
/dev_stage="Egg"
/tissue_type="oocytes"
1..3718
/standard_name="Xenopus poly(ADP-ribose) polymerase"
BASE COUNT 1105 a 798 c 954 g 861 t
ORIGIN

Query Match 33.2%; Score 157.6; DB 5; Length 3718;
Best Local Similarity 60.4%; Pred. No. 1.1e-38;
Matches 278; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

Qy 9 gctattggcagcaggttcaaggttgacgaatttgggaattcttagtcaagggttaag 68
Db 2657 GCTGTTGGCAGCGCTCCGACCCACGACCAATTTTGCAGGAATATTGCTCAGGCTCCG 2716

Qy 69 aattcacactcctcagcagcctgttactgtatgttcgcgaagccctctactttgc 128
Db 2717 AATTCTCCGCAAGAGCTCTGTACCGGGTATGTTTGGCAAGGTATCTATTTCG 2776

Qy 129 agatcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 188
```



```

Db 2777 CGCATGGTATCCAAAGTCAAACTACTGTCTGCAATGCGGGTAGCCCCATAGGGCT 2836
Qy 189 gatctcttctgaggttctttagagagacatgtatgaactaaagaagccagtcctat 248
Db 2837 GATCTTACTGGGGAGGTGGCCCTCGGAAACATGCAATGAGCTGAAAGCAGCATCAAAAT 2896
Qy 249 ggacaaacctccaagaggaagcattcgacacagggattaggcaaaacccgtgccactga 308
Db 2897 TACAAACTCCCAAGGCAACACAGTGAAGGGTTTGGGTAGAAGTGCACAGATCC 2956
Qy 309 gtcagagtttgaagtgagggagtgatgctagttccctgaggaagccggtgccatc 368
Db 2957 CTCAGTACTGTACA--GCTGGATGAGTGGATGTTCTCTCGGGAAGGAACATCTCGC 3013
Qy 369 atcaattaggagctgaactcatcatgaatgagtagcatctctcaacacacacccaggt 428
Db 3014 AAATATCAGTACACACCCCTGTTGTATATGAATATATTGTGTATGATATATTCTCAGGT 3073
Qy 429 gaagatgcagttcttctggaaggtgcgtttccatcacaag 468
Db 3074 CAACCTGAAGTACCTGCTGAAGTCAAGTTCACACTACAAG 3113

RESULT 6
LOCUS RNU94340 3128 bp mRNA ROD 19-FEB-1998
DEFINITION Rattus norvegicus poly(ADP-ribose) polymerase mRNA, complete cds.
ACCESSION U94340
VERSION U94340.1 GI:2896791
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE Benke S., Meyer R. and Buerkle A.
AUTHORS Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly
TITLE (ADP-ribose) polymerase
JOURNAL Biochem. Mol. Biol. Int. 43 (4), 755-761 (1997)
MEDLINE 98046546
REFERENCE Benke S., Meyer R. and Buerkle A.
AUTHORS Direct Submission
TITLE Submitted (18-MAR-1997) Angewandte Tumoriologie, Deutsches
JOURNAL Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
D-69120, Germany
REFERENCE Benke S., Meyer R. and Buerkle A.
AUTHORS Direct Submission
TITLE Submitted (19-FEB-1998) Angewandte Tumoriologie, Deutsches
JOURNAL Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg
D-69120, Germany
REMARK Sequence update by submitter
COMMENT On Feb 19, 1998 this sequence version replaced gi:2583143.
FEATURES
source 1. .3128
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_type="mononuclear blood cell"
31. .3075
/codon_start=1
/product="poly(ADP-ribose) polymerase"
/protein_id="AAC53544.1"
/db_xref="GI:2896792"
/translaton="MAEATERYLYRVEYAKSGRASCCKSESTPKSLRMAIMVQSPMF
DGKVPWHYHPCSCFWKVGHSIROPTEDVDFSELRWDDQKVKKTAAGVAGKGGHGG
GGAKEITLGDFAEYAKSNRSTKCGMEKIERGOMRLSKMLDPEKPOLGMDIRWHP
TCFVKNDELGFPEYASOLKGFSLSAEDKALKKOLPAVKSEGRKRCDEVDIDE
VAKKSKGKDKSESSKLEALKAKQNELVNIKDELKACSTNDLKELLIFNOQVPSG
ESAILDRVADGMAGFALLPCKECSGLVFKSDAYCTGDTAWTKMVKTNPSRKEW

```

BASE COUNT 861 a 758 c 905 g 604 t
ORIGIN

Query Match 33.1%; Score 157; DB 10; Length 3128;

Best Local Similarity 59.9%; Pred. No. 1.7e-38;

Matches 281; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```

Qy 1 aacaagatgctattatggcaggttcaaggttgacgaatttttg9ggaattcttagtcaa 60
Db 2596 RACCGAGACTGCTGTGGCAGCGGTCCAGGACCACCAACTTCGACGGCATCTCTGTCA 2655
Qy 61 gggctaagaattgaacctctgagcagcctgttactgcttatgtcttcgcaagagcc 120
Db 2656 GGTCTGCGGATAGCCGCCACCTGAAGCACCTGTGACAGGCTACATGTTGGGAAAGGA 2715
Qy 121 tactttgagatctagtaagaagcagcacaactactgttatgtgtagtagaataatcct 180
Db 2716 TACTTTGCTGATATGCTGTCCAAAGTGCAGAACTACTGCCACAGCTCTCAGGGAG 2775
Qy 181 gtaggtttgatgctcttctgaggttctttagagagacatgtatgaactaaagaagcc 240
Db 2776 ATTGGCTTAATAGTGTGGAGAGTGTCCCTTGGAAACATGTACGAACTCAACATGCT 2835
Qy 241 acgtccatgacaaacctccaagagggaagcattccacaaagggattaggcaaacccgtg 300
Db 2836 TCTCATACAGCAAGTATCCCAAGGCAAGCACAGTGTCAAAGGTTTGGGCAAAACCG-- 2893
Qy 301 ccaactgagtcagagttgtgagtgaggagatgatgtcgtagttccctcgaggcaagccg 360
Db 2894 -CCCTTGACCTTCGGCCAGCATCACCTCTGATGTTAGAGGTTCCCGCTGGGAACAGGG 2952
Qy 361 gtccatcatcaattaggagctcgaactcatgtacaatgagtagcatcgtctcaacaca 420
Db 2953 ATTCCCTCTGTTTAATGACACCTCCCTGCTGTAATACGAGTACATGTTCTACGACATT 3012
Qy 421 tcccaggtgaagatgcagttcttctgctgaaggtgcgtttctcatcacaaga 469
Db 3013 GCTCAGTGAATCTGAAGTACTACTGAAACTGAAGTTCATTTCAATTCAAGA 3061

```

RESULT 7

XPARP 3617 bp mRNA VRT 18-JAN-1993
LOCUS X laevis PARP gene encoding poly(ADP-ribose) polymerase.

DEFINITION

ACCESSION Z12139

VERSION Z12139.1 GI:64967

KEYWORDS poly(ADP-ribose) polymerase.

SOURCE African clawed frog.

ORGANISM Xenopus laevis

REFERENCE

AUTHORS

JOURNAL

de Rennes 1. URA CNRS 256

REFERENCE

AUTHORS

JOURNAL

Sautier-le Dreaan, B.M.

Direct Submission

TITLE

JOURNAL

Submitted (15-MAY-1992) Sautier-le Dreaan B.M., Lab. de Biol. et

Genet. du Développement, Université de Rennes I, URA CNRS 256,
Campus de Beaulieu, Av. du Gal LeClerc, Rennes-cedex, FRANCE, 35042
Location/Qualifiers
1. .3617
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/germline
/sex="Female"
/dev_stage="Adult"
/tissue_type="Ovary"
/cell_type="Oocyte"
/clone_lib="oligo dt primed lambda gt10 library of D.
MELTON"

FEATURES

source

gene
1. .2997
/gene="PARP"
CDS
1. .2997
/gene="PARP"
/EC_number="2.4.2.30"
/function="poly ADP-ribosylation, DNA repair"
/note="Protein sequence is in conflict with the conceptual
translation; poly(ADP-ribose) polymerase"
/citation=[1]
/codon_start=1
/product="NAD(+) ADP-ribosyltransferase"
/protein_id="CAA78126.1"
/db_xref="GI:1334661"
/translation="AKGRASCKKCGGNIKASISGLAIWQSMFQDGKVPWHHVSFCF
WKRARVLSQGDYICVTELWEDQEMIKAJETGGAAGGDSKGGKGMETLNDFAAE
YAKNSRSCGCEQKIEKQIRSKSVDERPOLGMDIRWHPDCFVSSREELDFLP
SYSASQLKFTLLSAEDKSDLSKRPVKNRGRKADVDGHSAAATKKIKKEKES
KLELLKEQTEILWHIKDELKVCSTNDLKELIANKQVPSGETNIVRSDGMAFG
ALLPECSQGFVKDAYCYCTGDSAWTKCAVKTOTPNRKDWVTPKEFETPYLKKF
FKRRHRAFPPTCAAPTISPAPAEKPTVEETPEGKPLTNTKVLGKLSKNDKDE
KTLIEGLGGVAGSAHKANLCISTNKEVKMKMEVEKAAVRVSDDFLKEVSEK
SVQELLQSGFTSSWGAETKOEAVQTEKOPSSGPVAGKSGKVKKEKSGKSEKMKML
TVKGAALDPSELEDSCHVLEGTGGIFSATLGLVDITRTNSYKQLQIIEHRSYPL
WFRPSHWRGTVIGSKLEMSKDEAIEHLNLYQDGTGNAMHSPNTKPKFEPYPL
EIDYQGEDVVKLSVAGQPKSLKAPQVELIKLIFDVESMKAMVEFIDIQKPLG
KLSKQIQSAYSLSQVQAVSELSSEARLLDSNOFYTLIPHDGFMKPPLLNLEY
IQAKVOMLNDLIEVAYSLIRGGADGGEKOPIDVYKIKITDIDKVAQDSFESPIIC
DYKNTHADPHNADLEVEIFIKIDREGEYQYKPKOLHNQOLLHSGSRNTFAGIL
SOGRIAPPAPVGYMGKGIYFADMVSKSANYCHAMPGSPIGLILGEVALGNHIE
LKAASQITKLPKCKHSGVGLGRTAPDPSAIVQDGVDPVPLGKGTSAISDTSLLYNEY
IVVDIAQVNLKLLKLFNKYKGMW"
3'UTR
2995. .3617
polyA_signal
3551. .3556
polyA_site
3574
BASE COUNT 1131 a 757 c 910 g 819 t
ORIGIN

Query Match 32.7%; Score 155; DB 5; Length 3617;
Best Local Similarity 60.1%; Pred. No. 7.3e-38;
Matches 276; Conservative 0; Mismatches 180; Indels 3; Gaps 1;
Qy 9 gctattatgacaggttcacgaattttgttggaattcttagtcaagggtcaag 68
Db 2573 GCTGCTTGGCAGCGGCTCCCGACCAACAAATTTTGAGAAATATTGTCTCAGGGTCTCGC 2582
Qy 69 aattgacacctctgagggcacctgtactgctatatgttcgcaaggcctcactttgc 128
Db 2583 AATTGCTCCGCCAGAGCTCCTGTTACCGGGTATATGTTTGCAAGGATATCTATTTTGC 2642
Qy 129 agacttagtaagcaagagcacaatactgtattgttgtaggaataataatcctgtagttt 188
Db 2643 CGACATGGTATCCAAAGTGCAAACTACTGTCTATCAATGTCGCGGTAGCCCATAGGCT 2702
Qy 189 gactgtctcttctgaggtgtcttaggacacatgtatgaactaaagaagcagctcoac 248
Db 2703 GATCTTACTGGGGAGGTGGCCCTCGGAACATGATGATGATGATGATGATGATGATGAT 2762

Qy 249 ggcacaaacctccaagaggaagcattcgacccaagggattaggaacaaacccgtgccactgga 308
Db 2763 TACAAAACCTTCCCAAGGCAACACACAGTGTAAAGGGTTTGGGTAGACTGACACAGATCC 2822
Qy 309 gtcagagtttgtgaagtggagggatgatcgtagtctccctcgcggaagccggtgcccac 368
Db 2823 CTGAGCTACTGTACA---GCTGGAGTGGATGATGTTCTCTCGGGAAGGAACCTTCGC 2879
Qy 369 atcaattaggagctctgaactcatgtacaatgacatgacatcgctcacaacacatcccaagt 428
Db 2880 AAATATCAGTGACACACAGCTGTTGTATATGAATATATTTGTATGATATTTGCTCAGGT 2939
Qy 429 gaagatgcagttcttctgctgaaggtgctgtccatcacaa 467
Db 2940 CAACCTGAAGTACTGCTGAGGCTCAAGTCAACTACAA 2978

RESULT 8

ATPARP 2147 bp mRNA PLN 31-MAY-1995
LOCUS A.thaliana PARP mRNA for PARP protein.
ACCESSION Z48243
VERSION Z48243.1 GI:853721

PARP gene: PARP protein.
thale cress.
Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 2147)
Lepiniec L., Babiychuk, E., Kushnir, S., Van Montagu, M. and Inze, D.

AUTHORS

Characterization of an Arabidopsis thaliana cDNA homologue to

TITLE

animal poly(ADP-ribose) polymerase

MEDLINE

FEBS Lett. 364 (2), 103-108 (1995)

REFERENCE

2 (bases 1 to 2147)
Lepiniec L.

AUTHORS

Direct Submission

TITLE

Submitted (10-FEB-1995) Lepiniec L., University Gent, Laboratorium

JOURNAL

Genetica, 35 KL Ledeganckstraat, GENT, Belgium, B-9000

FEATURES

Location/Qualifiers

1. .2147
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="adp-10"
1. .2147
1. .128
129. .2042
/gene="PARP"
129. .2042
/gene="PARP"
/standard_name="Poly(ADP-riboseyl) polymerase"
/EC_number="2.4.2.30"
/citation=[1]
/codon_start=1
/evidence=experimental
/product="PARP protein"
/protein_id="CAA8288.1"
/db_xref="GI:853722"

source

1. .2147

1. .128

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

129. .2042

CDS
1..3036
/note="poly (ADP-ribose)polymerase (AA 1-1011)"

py 1 aacaagatgctattatggcacggttcagggtgacgaattttgtgggaattcttagtcaa 60

RESULT	11
ZMPARP1	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2295)
Babyichuk E., Cottrell, P., Storozhenko, S., Fuangthong, M., O'Farrell, M., Van Montagu, M., Inze, D. and Kushnir, S.
Higher plants possess two poly(ADP-ribose) polymerases
2 (bases 1 to 2295)
Kushnir, S.
Direct Submission
Submitted (14-NOV-1997) Kushnir S., VIB, Dep. Genetics, Ledeganchstraat 35, Gent, B9000, Belgium
Location/Qualifiers
1. .2295
/organism="Zea mays"
107. .2068
/db_xref="taxon:4577"
/gene="PARP"
107. .2068
/gene="PARP"
/function="secondary protein modification"
/codon_start=1
/product="poly(ADP-ribose) polymerase"
/protein_id="CAA10888.1"
/db_xref="GI:2959360"
/db_xref="SPTREMBL:O50017"
/translation="MSARLRVADVRAELQRLGDLVSGTKPALVRRLDAAITCEAEKAAV
AAAPTSTVANGDVADCGKRCNCKNKKRSGDGGEGNGDCTDVTDKLEGMSYRELQGL
AKARGVAANGKGVDIQRLLSATAGPAADVGGPLIGAKEIVIKGGDEEVEYKKEKMWTA
TKKGAVALDQIHDPHTKYNHVLQGVDEIDYIATLNQTVNGDNNKFEYIIQVLESADGA
SEMYVNRGRVGRVQDQLKGGSPTRDQALIEFEFGKFNKTNHNSDRKNFKCYAKY

TWLEMDYGETEKEIEKGSITDQIKETKLETRIAQFISLICTISMKMKRMEVIGYNAEK
 LPLGKLKATILKGVHLKRIKSDVSKADRRHLEQLTGEFYVILPHDFCFKRMREFII
 DPOKLRKALVEALGIEIATILKLEDDSDDDPLIARYKOLHCDTFLPADSDEY
 SMKSLRTHKTHGTSYTDVIOIFKVSRRHGETERFQKFASTRMILHWSRLSNW
 AGTISGLRIAPPEAPVTVGKMGVYFADMFSAKSYCYASACKSVLLICEVALG
 DMNELLNADYDANNIPKGLSKSGVGTAPNWKSVKADVVVVVPLGPKQEPKSRGG
 LLYNEYIVYNVDIIRMYVHLVHNFERR"
 BASE COUNT 668 a 447 c 597 g 583 t
 ORIGIN

Query Match 31.9%; Score 151.4; DB 8; Length 2295;
 Best Local Similarity 59.0%; Pred. No. 9.3e-37;
 Matches 279; Conservative 0; Mismatches 191; Indels 3; Gaps 1;
 QY 1 acaagatgctattatggcaggttcgaaggtgacgaatttgggaattcttagtcaa 60
 Db 1589 AATAGGATGCTTTTGGGATGCTTCTCGTTTGGACCACTGGGCTGGATCTTCTCAG 1648
 QY 61 gggcctaagaattgcactctgaagcactgttactggtctatatgttcggcaagccctc 120
 Db 1649 GGTCTGGAATCGCTCTCTGAGCACCTGTTACTGTTTACATGTTTGGCAAGGGTGT 1708
 QY 121 taccttcagatctagtaagcaagcgcacaaactactattatgtaggaataatcct 180
 Db 1709 TACTTGTGACATGTTTCAAGAGAGTGCAAACTATGTACGCTCTGAAGCATGTAGA 1768
 QY 181 gttaggttgatgcttcttctgaggttctttaggagacatgtatgaactaaagaagcc 240
 Db 1769 TCTGAGTACTGCTTTTATGAGTGGATGCTTGGCGGATGATGAATGAGTCTGAATGCA 1828
 QY 241 acgttcattg---acaaacctccaagaggaagcattcgaaggaattaggaagcc 297
 Db 1829 GATTAGATGCTAATAACCTGCCAAAGAAATTAAGATCCAAAGGAGTGTGTCACAAAC 1888
 QY 298 gtgcactgagctagctgaggttctgaggtgaggtgaggtgaggtgaggtgaggtgag 357
 Db 1889 GCACCTTAACATGCTGAGTCTAAGGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 1948
 QY 358 ccggtgcatcatcaattaggagctctgaactcatgataatgagtagtactctcacaac 417
 Db 1949 CCCAAACAGGAACCTTCCAAAAGGGGGTGGCTGCTTTATATGATGATGATGATGAT 2008
 QY 418 acatccagatgaagatgcaggtcttctggaaggtgcgttccatcaagag 470
 Db 2009 GTAGACCAGATAGCAATGCGGTATGCTTACATGTAACTTCAATTTCAAGAG 2061

RESULT 12
 MMADPRP 3172 bp mRNA ROD 12-SEP-1993
 LOCUS Mouse mRNA for poly (ADP-ribose) polymerase (EC 2.4.2.30).
 DEFINITION X14206
 ACCESSION X14206.1 GI:49893
 VERSION DNA-binding protein; NAD(+) ADP-ribosyltransferase; zinc finger protein.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3172)
 Huppi, K.
 Direct Submission
 Submitted (26-JAN-1989) Huppi K., NCI/NIH, Lab of Genetics, Building 37, Room 2B-21, Bethesda, MD 20892
 2 (bases 1 to 3172)
 Evers, R., Hammer, A. and Cornelissen, A.W.C.A.
 Unusual C-terminal domain of the largest subunit of RNA polymerase II of Chithidia fasciculata
 II of Chithidia fasciculata
 Nucleic Acids Res. 17, 3387-3401 (1989)
 Location/Qualifiers
 1..3172
 /organism="Mus musculus"

/strain="BXS8"
 /db_xref="taxon:10090"
 /tissue_type="spleen"
 /clone="pADP3-2.1 and pADP-3C"
 85, 3126
 /note="poly (ADP-ribose) polymerase (AA 1 - 1013)"
 /codon_start=1
 /protein_id="CAA32421.1"
 /db_xref="GI:49894"
 /db_xref="MGD:MGI-87933"
 /db_xref="SWISS-PROT:P11103"
 /translation="MAEASERLYVOYAKSGRACKSCSEIPKDSLRMAIVQSPMF
 DGKPHWHYHSCFWKQVGSIRHPDVEVDESELRWDDQKVKTAAGAGVAGKQDGS
 GKAETKLGDAFAEYAKSMKCGCKLEIKKGMRLSKKMWDPPEKQMGMDRWVHP
 TCVKRDELGFPEYSASQLKGLSAEDKELKQKPAIKNEKRGKRGDEVDGDE
 VAKSKETDKYKLEKAKAQNELIWNIDELKACSTNDLKEILLIFNOQOVPSGE
 SAIDRVADMGAFGLPKCEGSLQVFSAYICTGDTAVTAVNNSAPADKPLSN
 TPREFREISLYLKLKVKYKODRIFPPPESSAPITVHPLSVTSAPTAVNSAPADKPLSN
 MKILTLKSONKDEAKVTEKLGKLTGSANKASLCTISIKKEKMKMKWEVKKAN
 IRVYSELDQVDSASTKSLQDLKLSAHSLSMPGAHVLEKGVKVFSAITGLDVIKGTN
 CFKEGVNKESEKRMKLTGKGAAYDPDGLSHSAHLEKGVKVFSAITGLDVIKGTN
 SYKLOLLEDDKESRYWIFKSWRGITVTSNKLQMPKSKLPKQVLEWIMFIDVDSMK
 WHSKNFTYKPKFPLEIDYQDEEAVKALTVKPGTCKSLKPKQVLEWIMFIDVDSMK
 KALVEYEDLQKPLGLSRQIOAAYSILSEVQVPSGSSQSDSKDPIDVNYEKLTP
 HFGMKPPLNNADSVQAKVEMLDNLDIEVAYSLRGSSDSDSKDPIDVNYEKLTP
 DTKVDRDSBEAEVIRKTVKNTHTTHNLDIEVAYSLRGSSDSDSKDPIDVNYEKLTP
 RLWHGSRITNFAGILSQGLRIAPPEAPVTVGKMGVYFADMFSAKSYCYASACKSVLLICEVALG
 IGLIMGEVALGNMTELKHAHSHISKLPKHSYKGLGKTPDPSPASITLEGVEVPLGT
 GIPSGVNDTALLYNIYVIDIAQVNLKYLKLEKNEFTSLW"
 misc_feature 88..678
 /note="zinc finger domains"
 misc_feature 2743..2823
 /note="DNA-binding domain A"
 misc_feature 2896..2952
 /note="DNA-binding domain B"
 /note="DNA-binding domain B"
 BASE COUNT 869 a 779 c 925 g 599 t
 ORIGIN

Query Match 31.8%; Score 150.6; DB 10; Length 3172;
 Best Local Similarity 59.1%; Pred. No. 1.8e-36;
 Matches 277; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
 QY 1 acaagatgctattatggcaggttcgaaggtgacgaatttgggaattcttagtcaa 60
 Db 2647 AACCGGAGGCTGCTGTCGCGCCGCTCCAGGACCAACCTTGTGCGGACGCTGCTG 2706
 QY 61 gggcctaagaattgcactctgaggaacacctgtactgctatatgttcggcaagccctc 120
 Db 2707 GGTCTGCGATAGCCACCTGAAGCGCCGCTGACAGGCTACATGTTTGGGAAAGGATC 2766
 QY 121 taccttcagatctagtaagcaagcgcacaaactactgttatgttgataggaataatcct 180
 Db 2767 TACTTGTGCGATGCTGTCGCGCCGCTCCAGGACCAACCTTGTGCGGACGCTGCTG 2826
 QY 181 gttaggttgatgcttcttctgaggttctttaggagacatgtatgaactaaagaagcc 240
 Db 2827 ATTGCTTTAATAATGCTGGGAGGTTGCCCTTGGAAATATGTATGAACCTCAAGCATGCT 2886
 QY 241 acgttcattgcaaacacctccaagaggaagcattcgaacccagggattaggaagccg 300
 Db 2887 TCACATATACGCAAGTTTACCCCAAGGCAAGCAGAGTGTCCAGAGTTCCACTGGGAAACCG 360
 QY 301 ccaactggagtcagagttgtgaaagtggaggtgaggtgaggtgaggtgaggtgaggtgag 303
 Db 2944 ACCCTTGACCTTCGCGCCGCTGACATCACCCTGAGGGGTGTAGAGGTTCCTACTGGGAAACCG 420
 QY 361 gtgcactcatcaattaggagctctgaactcatgtacacatgacatcgctctcacaacaca 3063
 Db 3004 ATCCCACTCTGGTCTCAACGACACTGCCCTGCTGTATATAGTACATGCTGTACGACAT 421
 QY 421 tccagatgaagatgcaggttcttctggaaggtgcgttccatcacaaga 469

Db 3064 GCTCAGGTGAATCTCAATACCTGCTGAACTCAAGTTCAATTTAAGA 3112

RESULT 13

BOVPAS 3589 bp mRNA MAM 07-FEB-1999

DEFINITION Bovine poly(ADP-ribose) synthetase (EC 2.4.2.30) mRNA.

ACCESSION D90073

VERSION D90073.1 GI:217581

KEYWORDS DNA binding protein; helix-turn-helix; nuclear location signal; poly(ADP-ribose) synthetase; zinc-binding finger motif.

SOURCE Bovine thymus, cDNA to mRNA, clone p50-7.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bos (bases 1 to 3589)

REFERENCE Saito, I., Hatakeyama, K., Kido, T., Ohkubo, H., Nakanishi, S. and Ueda, K.

AUTHORS Cloning of a full-length cDNA encoding bovine thymus poly(ADP-ribose) synthetase: evolutionally conserved segments and their potential functions

TITLE Gene (1990) In press

JOURNAL These data kindly submitted in computer readable form by: Isao Saito

COMMENT Department of Clinical Science and Laboratory Medicine Faculty of Medicine, Kyoto University

Shogoinkawahara-cho 54, Sakyo-ku

Kyoto 606

Japan

Phone: 075-751-3467

Fax: 075-771-4792

Zinc-binding finger motifs are observed in seq. 21-51 and 128-165. The sequences, 200-220 and 250-270, showed helix-turn-helix structure. Nucleotide-binding fold was found in seq. 890-903.

FEATURES

source Location/Qualifiers

1..3589

/organism="Bos taurus"

/db_xref="taxon:9913"

77..3127

/note="poly(ADP-ribose) synthetase"

/codon_start=1

/protein_id="BA014114.1"

/db_xref="GI:217582"

/translation="MARSSDKLRYEYKSGRACKKESIPKDSIRMAFWESPMP DKGIPHWLSCFWKVGSIWHPDVEVEGSELRWDDQTIKMAETGGRDVSCKGO YHPKCFVORREELGFEPSATHLMGFSVLTAEQDQTLKQVLPKPOLGMVDCW IDEVTKKSKKKEIKLEKALKAOQNLINWVDELKACSTNDLKELLIFNKEVVP SGESAILDRVADGVFGALLPCECSQGLVFKGDYCTGDTAWTKMVKQTTPNPK EMWTPKFEISYFKLKKIKKODRIPEPSPSPVGAAPSAAPAAVHSGPPDKPL SNMKILTLGKLSQNDKVKATEKLGKLTGTANKASLCISTKKEVDKLNKWEVKE ANIRVSEDFQDISASTKSLQSLGSLHLLSPWAEVKEVPEAVGPKGSGRAPSCK SKGPVKEEPTNKEKRMKUTLAGGAADPDGSLHNAHLEKGGKVPFSAVLGLVILK GTNSYKQLLEDKESRYWIFRSRGVTGVSGLKLEOMPSKEDALEHPKLIYEET GNAAHKSFTFKPKKPYLEIDYQDEEAVKLTVPNCTKSLKLPVQNLKMFIDEV SMKAMVEYIDQLQMKPLKSKQIOAAYLSILSEVOALQSQSSSHLDLSNRYET LIHPDKMKPPLNNANSVQAVEMLDLIDEVATSLRGGSDSSSDPDIDVNEYK HNRRLWHGSRITNAGILSGLRIAPPEAPVTGMYFGKGIYFADVMVKSANVCHTSO GDPGILGLLEAGLNMVYELKHARHRIAPPEAPVTGMYFGKGIYFADVMVKSANVCHTSO LGTGISGVNDPCLLYNEYIYDIAQVHLKLLKLFNFKTSLW"

misc_feature 224..231

/note="nuclear location signal"

BASE COUNT 975 a 879 c 1010 g 725 t

ORIGIN

Query Match 31.8%; Score 150.6; DB 4; Length 3589;

Best Local Similarity 59.1%; Pred. No. 1.8e-36;

Matches 277; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

Qy 1 acaagatgctattgacaggttcaaggttgacgaattttgtggaattcttagtcaa 60

Db 2648 AACCGAGGCTCTGTGGCAGCGGTCCAGGACCAACCACTTCGGGGCATCTGTCCAG 2707

QY 61 gggttaagaattgacacctctctgagggacacctgttaactggttatatgttcggaagggctc 120

Db 2708 GGTCTCCGGATAGCCACCTGAGGACCTGTGACGGGCTACATGTTTGGTAAGGAATC 2767

QY 121 tactttcagatctagtaagaagagcacaatactgtttatgtgtgataagaaatcct 180

Db 2768 TATTTCGGGACATGGTCTCCAGAGTGCCTTGGGAGAGCTGCCCTTGGAAACATGTATGAATTAACATGCT 2887

QY 181 gtatgtttgatgtcttcttctgaggttgccttttagagagacatgtatgaactaaagaagcc 240

Db 2828 ATAGGCTTGATCTGTGGGAGAGCTGCCCTTGGAAACATGTATGAATTAACATGCT 2887

QY 241 acgtccatgagacaaacctccaagagggagcattcgcacacgctccagggagaccca 2827

Db 2888 CGACATATCAGCAAGTATCCCAAGGGCAAGCACAGTGTCAAAGGTTTAGGCAAAAC--T 2944

QY 301 ccaatgagtcagagttgtgaagtggagggatgatctatccctgcgcgaagccg 360

Db 2945 ACCCTGACCCATCGCTAGTATTACTGTGATGGTGGAGTGGCTCTCGGAGACGGG 3004

QY 361 gtgcatcatcaattagagctctgaactcatgtacaaatgagtcacatcgtctacacaca 420

Db 3005 ATTTTCATCTGGTGTATGACACCTGTCTGTGTATAGGAGTACATCGTCTACGACATT 3064

QY 421 tcccaagtggaagtcagttcttctgtaaggtgcgtttccatcacaaga 469

Db 3065 GCTCAGGTCCACCTGAAAGTACCTGCTGAAAGTCAAGTTCACTTTAAGA 3113

RESULT 14

AF126717 1732 bp mRNA ROD 16-MAY-2000

LOCUS AF126717

DEFINITION Mus musculus msparp mRNA, complete cds.

ACCESSION AF126717

VERSION AF126717.1 GI:7363321

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 1732)

AUTHORS Sallmann, F.R., Vodenicharov, M.D., Wang, Z.Q. and Poirier, G.G.

TITLE Characterization of SPARP-1. An alternative product of PARP-1 gene with poly(ADP-ribose) polymerase activity independent of DNA strand breaks

JOURNAL J. Biol. Chem. 275 (20), 15504-15511 (2000)

MEDLINE 20270268

REFERENCE 2 (bases 1 to 1732)

AUTHORS Sallmann, F.R., Vodenicharov, M.D., Wang, Z.-Q. and Poirier, G.G.

TITLE Direct Submission

JOURNAL Submitted (08-FEB-1999) Health and Environment Unit, Laval University Medical Research Center, CHUQ and Faculty of Medicine, Laval University, 2705, Blvd Laurier, Ste-Foy, Quebec G1V 4G2, Canada

FEATURES

source Location/Qualifiers

1..1732

/organism="Mus musculus"

/strain="129/Sv x C57BL/6"

/db_xref="taxon:10090"

/cell_type="fibroblast"

/dev_stage="embryo"

1..251

252..1730

/codon_start=1

/product="msparp"

/protein_id="AAF61293.1"

/db_xref="GI:7363322"

5'UTR

CDS

KKALVEYEDLDKPLGKLSRROIAAAYSLSEVQEPVSGSSSESOILDLSNRFTLLI
PHDFGKKPPLNNADSVLAKVEMLDNLLDIEVAYSLRGSDSSDKDPIDVNYEKLK
TDIKVDRDSEAEVIRKYVKNTHATTHNADVIDEIDFKIEREGESORYAPFROLHN
RLLWHGSRFTNAGILSOGRIAPPEAPVTVGFMGKGIYFADVMVKSANTYCHTSQGD
PIGLJMLGNEELKNNYELKSHASKPKGKHSYKGLKTKTPDPFSASITLSEVVEPLG
TGIPSGVNDTCLLYNEYIVYDIAQVNLKYLKLFNFKTSLW

3'UTR
BASE COUNT 472 a 424 c 493 g 343 t
ORIGIN

Query Match 31.4%; Score 149; DB 10; Length 1732;
Best Local Similarity 58.8%; Pred. No. 5.1e-36;
Matches 276; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

Qy 1 aacaagatgctattatggcagcgttcaaggttgacgaatttttgggaatttttagtcaa 60
Db 1251 AACCGGAGGCTGTGTGGCAGCGCTTCAGGACCAACCACTTGTGGCATCTGTGCGAG 1310
Qy 61 gggctaagaattgcactctgaagcagcctgttactgcttatgttgcgcaagccctc 120
Db 1311 GGTCTCGGATACGCCACCTGAAGCGCTGTGACAGGCTACATGTTTGGGAAAGGATC 1370
Qy 121 tactttgcagatctagtaagcagagcgacacatactgttattgtgtaggaataatcct 180
Db 1371 TACTTTGCCACATGTTGCCAAGTCAAACTACTGCCACACATCTCAGGGAGACCCG 1430
Qy 181 gtagggttgatgcttcttctgtgaggttcttttaggagacatgtatgaactaagaagcc 240
Db 1431 ATTGGCTTAATAATGCTGGGAGAGGTTGCCCTTGGAAATATGATGAACATCAAGCATGT 1490
Qy 241 acgtccatgacaaacctccaagaggaagcattcgaccaaaggattaggaacaaaccgtg 300
Db 1491 TCACATATCAGCAAGTTACCAAGGCAAGCAGACAGTGTCAAGGTTTGGGAAAAACC --- 1547
Qy 301 ccactgagtcagagttgtgaagtggaggtgatgtgtgtagtccctgcggcaagccg 360
Db 1548 ACCCTGACCTTCGGCCAGCATCACCTGGAGGGTGTAGAGGTTCCTACTGGGAACAGGG 1607
Qy 361 gtgccatcataataggagctctgaactcatatgacaaatgacatgacatgcttacaacaca 420
Db 1608 ATCCATCTGGTGTCAACAGACACCTGCTGCTGTATTAATAGTACATGTCGTACGACATT 1667
Qy 421 tccagtggaagatgcagttcttctgtaaggtgcgttttccatcacaca 469
Db 1668 GCTCAGGTGAATCTCAATACCTGCTGAACACTCAAGTTCATTTAAGA 1716

RESULT 15
AF168781 3047 bp mRNA ROD 28-JUL-1999
LOCUS
DEFINITION
Cricetulus griseus poly ADP-ribose polymerase (PARP) mRNA, complete cds.
ACCESSION
AF168781
VERSION
AF168781.1 GI:5616519
KEYWORDS
Chinese hamster.
SOURCE
Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.
REFERENCE
1 (bases 1 to 3047)
Ganesh,A.N., Phillips,E., Shah,R., Affar,E.B., Poirer,G., Thacker,J. and Meuth,M.
Suppression of the radiation sensitive phenotype of hamster irs1 and irs2 strains selected for resistance to 3 aminobenzamide Unpublished
2 (bases 1 to 3047)
Ganesh,A.N. and Meuth,M.
Direct Submission
AUTHORS
Submitted (14-JUL-1999) Oncological Sciences, University of Utah, 15 N. 2030 E., Salt Lake City, UT 84112, USA
JOURNAL
TITLE
Location/Qualifiers
FEATURES

source
gene
CDS

1. .3047
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/cell_line="V79-4"
1. .3047
/gene="PARP"
1. .3042
/gene="PARP"
/EC_number="2.4.2.30"
/codon_start=1
/product="poly ADP-ribose polymerase"
/protein_id="A4D5817.1"
/translation="MAEASRLRYVEYAKSGRACKSCSESIPKDSLRMAIMVQSPMF
DGKVPHWHFSCFWKVGHSIRQPDVEVDGSELRWDDQQRVKTATAGGAGGAGGDS
GGSEKTLGDFAAEYAKNSSTCKGCMKEIKGOVRLSKMKMLDPEKPOLGMDIDWYHP
TCFVKNREELGFRPEYSASOLKGLSLLSAEDKEVLKQLPVKSEKGRKGDVDEGADE
VARKSKKGDKSKLEKALKANDLNIWNLDELKACSTSDLEKLLFNQOQVPSGE
SAILDRVADGMAGFALLPKECQQLVFKSDAYICTGDTAWTKMVKSTQTPSRKEWY
TPKPREISYLLKLVKKQDRIFPETSAPAPPLPSPVTSAPTAKMKMEVQAOAN
MKILTLGLKSQSDAEATKIEKGLKLTGSANNAISLCISTKKEVKMGKMAEVEQAN
VRVCEDFLODVAATKSLQELLASHLSWGAEVKEVPEVAAPKPKSAAPSKSGK
LYKEEGYNSEKMKLTLKGAADVDPDSGLEHSAHLEKGGKVFSAATGLVDIVKGTN
SYKLQLEDDEKESRYMIFRSWGRVGVIGSNKLEQMPSEDAVEHEHMKLYEKTGNA
WHSKNFTYPKKFIWIDYQDEEAVKLTVPKTKLKPKAQVLEHGMIFDVESHK
KALVEYEDLDKPLGKLSRROIAAAYSLSEVQAOVSGSSDSOILDLNRRFTLLI
HDFGKKPPLNNADSVLAKVEMLDNLLDIEVAYSLRGSDSSDKDPIDVNYEKLK
TDIKVDRDSEAEVIRKYVKNTHATTHNADVIDEIDFKIEREGESORYAPFROLHN
RLLWHGSRFTNAGILSOGRIAPPEAPVTVGFMGKGIYFADVMVKSANTYCHTSQGD
PIGLJMLGNEELKNNYELKSHASKPKGKHSYKGLKTKTPDPFSASITLSEVVEPLG
TGIPSGVNDTCLLYNEYIVYDIAQVNLKYLKLFNFKTSLW"

BASE COUNT 846 a 739 c 869 g 593 t
ORIGIN

Query Match 31.1%; Score 147.4; DB 10; Length 3047;
Best Local Similarity 58.6%; Pred. No. 1.8e-35;
Matches 275; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

Qy 1 aacaagatgctattatggcagcgttcaaggttgacgaatttttgggaatttttagtcaa 60
Db 2563 ATCCGAGGTTACTGTGGCAGCGGTCCAGGACCACTTGTGTGGCATCTGTGTCACAG 2622
Qy 61 gggctaagaattgcactctgagcagcctgttactgcttatgttgcgcaagccctc 120
Db 2623 GGTCTCGGATAGCCCACTGAAGCAGCCTGTGACAGGCTACATGTTTGGAAAAGGATC 2682
Qy 121 tactttgcagatctagtaagcagagcgacacatactgttattgtgtaggaataatcct 180
Db 2683 TATTTTGTGATGATGTGTCAAAAGTGCACACTACTGCCATACATCTCAGGAGACCCG 2742
Qy 181 gtagggttgatgcttcttctgaggttgcgttttaggagacatgtatgaactaagaagcc 240
Db 2743 ATTTGGTTCATCTGTGGGAGAGTGCCTTGGAAACATGTACGAACTCAAGCATGCG 2802
Qy 241 acgtccatgacaaacctccaagaggaagcattcgacaaaggattaggaacaaaccgtg 300
Db 2803 TCACATATCAGCAAGTTACCAAGGCAAGCAGACAGTGTCAAAAGGTTGGGCAAAAC ---G 2859
Qy 301 ccaactgagtcagagtttgtgaagtgagggatgatgtcgttagtccctgcggcaagccg 360
Db 2860 ACCCTTGACCCCTTCGGCCAGCATCACCTGAGGGTGTAGAGGTTCCACTGGGAGGAGG 2919
Qy 361 gtgccatcataataggagctctgaactcatgacaaatgacatgacatgcttacaacaca 420
Db 2920 ATCCCATCGGGTGTAAATGACACCTGCTTACTATATAACGAGGTACATCTGCTATGACATT 2979
Qy 421 tccagtggaagatgcagttcttctgtaaggtgcgttttccatcacaca 469
Db 2980 GCTCAGGTGAACCTGAATACCTGCTGAAACTGAAGTTTAATTTTAAGA 3028

Fri Mar 8 08:32:02 2002

Search completed: March 7, 2002, 16:49:01
Job time: 8171 sec

us-09-236-995d-3.rge

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 18:16:17 ; Search time 298.35 Seconds
(without alignments)
1362.065 Million cell updates/sec

Title: US-09-236-995D-3

Perfect score: 474

Sequence: 1 acaagatgctattatggca.....gtttccatcacagaggttag 474

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SID22/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	474	100.0	2949	20 AAX89542	Maize poly ADP-rib
2	474	100.0	3211	21 AAZ60615	DNA encoding the p
3	474	100.0	3212	21 AAZ60618	DNA encoding the p
4	154.4	32.6	2147	21 AAZ60617	DNA encoding the p
5	151.4	31.9	2295	21 AAZ60616	DNA encoding the p
6	142.6	30.1	3580	20 AAZ23799	Human poly(ADP-rib
7	142.6	30.1	3793	27 AAT13732	DNA-binding domain
8	139.4	29.4	3045	22 AAF63954	Human tankyrase2 r
9	137.8	29.1	5345	16 AAT04221	5.3 kb EcoRI fragm
10	136.2	28.7	2682	16 AAT04219	Poly(ADP-ribose)-p
11	122.6	25.9	864	22 AAC85320	3' end fragment of

Human poly(ADP-rib
hparp2 cDNA. Homo
Human brain PARP2
Human brain poly-A
Human ORF ORF2673
Fusion protein PAR
Human cDNA sequenc
Human colon cancer
Poly(ADP-ribose)-p
Enterococcus faeca
Human brain PARP3
Human uterus type
Probe #9163 used t
NAD(+) ADP-ribosyl
EST AA568817, frag
Colon tumour relat
H. pylori GHP0 150
Human cDNA clone (
Human secreted pro
Aspergillus oryzae
Codon-optimised mu
Shrimp white spot
Human cervical can
Arabidopsis thalia
ACNPV ORF 54, resi
Drosophila morphog
Drosophila morphog
Drosophila morphog
ACNPV genomic DNA
Human gene signatu
Human secreted pro
Human secreted pro
Codon-optimised HP
Stromelysin gene c

ALIGNMENTS

RESULT 1

AAX89542

ID AAX89542 standard; cDNA; 2949 BP.

XX AC AAX89542;

DT 06-OCT-1999 (first entry)

XX DE Maize poly ADP-ribose polymerase gene.

XX KW PARP; poly ADP-ribose polymerase; metabolism; maize; transform; ss.

XX OS Zea mays.

XX FH Key Location/Qualifiers

XX CDS 1..2949

FT /*tag= a

FT /product= "PARP"

FT /note= "Poly ADP-ribose polymerase"

FT /transl_except= (pos:1453..1455, aa:Xaa)

FT /transl_except= (pos:1495..1497, aa:Xaa)

FT /transl_except= (pos:1504..1506, aa:Xaa)

FT /transl_except= (pos:1534..1536, aa:Xaa)

FT /transl_except= (pos:1537..1539, aa:Xaa)

FT /transl_except= (pos:1561..1563, aa:Xaa)

FT /transl_except= (pos:1582..1584, aa:Xaa)

FT /transl_except= (pos:1588..1590, aa:Xaa)

FT /transl_except= (pos:1594..1596, aa:Xaa)

FT /transl_except= (pos:1975..1977, aa:Xaa)

FT /transl_except= (pos:2011..2013, aa:Xaa)

FT /transl_except= (pos:2014..2016, aa:Xaa)

FT /transl_except= (pos:2026..2028, aa:Xaa)

FT /transl_except= (pos:2047..2049, aa:Xaa)

FT /transl_except= (pos:2050..2052, aa:Xaa)

Db	2776	ccactggagtcgaggttttgaagtggaggatgatgctgtagttccctgcggaagcgcg	2833
Qy	361	gtgccatcatcaattaggaagctctgaactcatgtacaatgagtacatcgctctacaacaca	420
Db	2836	gtgccatcatcaattaggaagctctgaactcatgtacaatgagtacatcgctctacaacaca	2895
Qy	421	tcccaagtgaaagatgcagttcttctgctgaagtgcggttccatcacaaagaggtag	474
Db	2896	tcccaagtgaaagatgcagttcttctgctgaagtgcggttccatcacaaagaggtag	2949
RESULT 2			
AAZ60615			
ID	AAZ60615 standard; DNA; 3211 BP.		
XX	XX		
AC	AAZ60615;		
XX	XX		
DT	16-MAY-2000 (first entry)		
XX	XX		
DE	DNA encoding the poly(ADP-ribose) polymerase ZAP1 protein of Zea mays.		
XX	XX		
KW	ZAP1; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase; programmed cell death; apoptosis; growth rate; stress; cold; pathogen; pest; drought; heat; fungi; nematode; seed-shatter; ss.		
KW	KW		
KW	KW		
XX	XX		
OS	Zea mays.		
XX	XX		
FH	Key		
FT	CDS		
FT	113..3022		
FT	/*tag= a		
FT	/*product= "ZAP1 protein"		
XX	XX		
PN	W02000004173-A1.		
XX	XX		
PD	27-JAN-2000.		
XX	XX		
PF	12-JUL-1999; 99WO-EP04940.		
XX	XX		
PR	17-JUL-1998; 98US-0118276.		
XX	XX		
PA	(PLBZ) PLANT GENETIC SYSTEMS NV.		
XX	XX		
PI	Babiychuk E, Kushnir S, De Block M;		
XX	XX		
WP	WPI; 2000-182436/16.		
DR	P-PSDB; AAV68833.		
DR	XX		
XX	XX		
PT	Modulating cell death, growth and stress resistance in eukaryotes, specifically plants, used, e.g. to impart fungus or nematode resistance		
PT	-		
PT	XX		
XX	XX		
PS	Example 1; Page 79-84; 126pp; English.		
XX	XX		
CC	The present sequence encodes the ZAP1 protein of Zea mays. This protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed cell death or apoptosis, and is a nuclear enzyme. The ZAP1 polynucleotide sequences can be used for modulation of programmed cell death in eukaryotic cells. The method is used, specifically in plants, to induce or protect against, programmed cell death, depending on the extent to which PARP activity is reduced. Reducing expression of endogenous ZAP class PARP only is also used to modulate programmed cell death, to increase growth rate and to produce plant cells that are more tolerant of stress (cold, chemical treatments, pathogens, pests, drought, heat, etc., or during transformation). Particular applications are generation of plants that are resistant to fungi or nematodes; are male or female sterile; or have better seed-shatter properties. The methods are also used to improve growth of transformed plant cells (and derived calli or complete plants).		
CC	XX		
SQ	Sequence 3211 BP; 968 A; 604 C; 813 G; 826 T; 0 other;		

Query Match 100.0%; Score 474; DB 21; Length 3211;
 Best Local Similarity 100.0%; Pred. No. 1.3e-148;
 Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacaagatgctattatgacaggttcaaggttgacgaatttgggaattcttagtcaa 60
 Db 2549 aacaagatgctattatgacaggttcaaggttgacgaatttgggaattcttagtcaa 2608
 QY 61 gggctaagaattgacactctcagggcaccctgttactggctatatgttcggcaaggcctc 120
 Db 2609 gggctaagaattgacactctcagggcaccctgttactggctatatgttcggcaaggcctc 2668
 QY 121 tactttgcagatctagtaagaagcagcgcaacaatactgtttatgttgataggaataatcct 180
 Db 2669 tactttgcagatctagtaagaagcagcgcaacaatactgtttatgttgataggaataatcct 2728
 QY 181 gtaggttgatgctcttcttctgaggttgccttttaggagacatgtatgaactaaagaagcc 240
 Db 2729 gtaggttgatgctcttcttctgaggttgccttttaggagacatgtatgaactaaagaagcc 2788
 QY 241 acgtccatggacaaacctccaaagggaagcattccgaccagggttaggcaaaaccgtg 300
 Db 2789 acgtccatggacaaacctccaaagggaagcattccgaccagggttaggcaaaaccgtg 2848
 QY 301 ccactggagtcagaggttggtaagtggagggatgatgtcgttagtccctcgccaagccg 360
 Db 2849 ccactggagtcagaggttggtaagtggagggatgatgtcgttagtccctcgccaagccg 2908
 QY 361 gtgccatcaataataggagcttgaactcatgtacaaatgacatcgctctacaacaca 420
 Db 2909 gtgccatcaataataggagcttgaactcatgtacaaatgacatcgctctacaacaca 2968
 QY 421 tcccagggtgaagatgcagttcttctgctgaagtgctgttccatcacaaagaggtag 474
 Db 2969 tcccagggtgaagatgcagttcttctgctgaagtgctgttccatcacaaagaggtag 3022

RESULT 3
 AAZ60618
 ID AAZ60618 standard; DNA; 3212 BP.
 XX
 AC AAZ60618;
 XX
 DT 16-MAY-2000 (first entry)
 XX
 DE DNA encoding the poly(ADP-ribose) polymerase ZAP2 protein of Zea mays.
 XX
 KW ZAP2; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
 KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT CDS 81..3023
 FT /*tag= a
 FT /*product= "ZAP2 protein"
 XX
 PN WO200004173-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 12-JUL-1999; 99WO-EP04940.
 XX
 PR 17-JUL-1998; 98US-0118276.
 XX
 PA (PLB2) PLANT GENETIC SYSTEMS NV.
 XX
 PI Babiychuk E, Kushnir S, De Block M;
 XX
 DR WPI; 2000-182436/16.
 DR P-PSDB; AAY68839.
 XX

PT Modulating cell death, growth and stress resistance in eukaryotes,
 PT specifically plants, used, e.g. to impart fungus or nematode resistance
 PT
 XX
 PS Disclosure; Page 103-108; 126pp; English.
 XX
 CC The present sequence encodes the ZAP2 protein of Zea mays. This protein
 CC is a poly(ADP-ribose) polymerase (PARP) protein (also known as
 CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed
 CC cell death or apoptosis, and is a nuclear enzyme. The ZAP2 polynucleotide
 CC sequences can be used for modulation of programmed cell death in
 CC eukaryotic cells. The method is used, specifically in plants, to induce,
 CC or protect against, programmed cell death, depending on the extent to
 CC which PARP activity is reduced. Reducing expression of endogenous ZAP
 CC class PARP only is also used to modulate programmed cell death, to
 CC increase growth rate and to produce plant cells that are more tolerant
 CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,
 CC etc., or during transformation). Particular applications are generation
 CC of plants that are resistant to fungi or nematodes; are male or female
 CC sterile; or have better seed-shatter properties. The methods are also
 CC used to improve growth of transformed plant cells (and derived calli or
 CC complete plants).
 XX
 SQ Sequence 3212 BP; 974 A; 600 C; 815 G; 823 T; 0 other;

Query Match 100.0%; Score 474; DB 21; Length 3212;
 Best Local Similarity 100.0%; Pred. No. 1.3e-148;
 Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacaagatgctattatgacaggttcaaggttgacgaatttgggaattcttagtcaa 60
 Db 2550 aacaagatgctattatgacaggttcaaggttgacgaatttgggaattcttagtcaa 2609
 QY 61 gggctaagaattgacactctcagggcaccctgttactggctatatgttcggcaaggcctc 120
 Db 2610 gggctaagaattgacactctcagggcaccctgttactggctatatgttcggcaaggcctc 2669
 QY 121 tactttgcagatctagtaagaagcagcgcaacaatactgtttatgttgataggaataatcct 180
 Db 2670 tactttgcagatctagtaagaagcagcgcaacaatactgtttatgttgataggaataatcct 2729
 QY 181 gtaggttgatgctcttcttctgaggttgccttttaggagacatgtatgaactaaagaagcc 240
 Db 2730 gtaggttgatgctcttcttctgaggttgccttttaggagacatgtatgaactaaagaagcc 2789
 QY 241 acgtccatggacaaacctccaaagggaagcattccgaccagggttaggcaaaaccgtg 300
 Db 2790 acgtccatggacaaacctccaaagggaagcattccgaccagggttaggcaaaaccgtg 2849
 QY 301 ccactggagtcagagtttggtaagtggagggatgatgtcgttagtccctcgccaagccg 360
 Db 2850 ccactggagtcagagtttggtaagtggagggatgatgtcgttagtccctcgccaagccg 2909
 QY 361 gtgccatcaataataggagctcctgaactcatgtacaaatgacatcgctctacaacaca 420
 Db 2910 gtgccatcaataataggagctcctgaactcatgtacaaatgacatcgctctacaacaca 2969
 QY 421 tcccagggtgaagatgcagttcttctgctgaagtgctgttccatcacaaagaggtag 474
 Db 2970 tcccagggtgaagatgcagttcttctgctgaagtgctgttccatcacaaagaggtag 3023

RESULT 4
 AAZ60617
 ID AAZ60617 standard; DNA; 2147 BP.
 XX
 AC AAZ60617;
 XX
 DT 16-MAY-2000 (first entry)
 XX
 DE DNA encoding the poly(ADP-ribose) polymerase NAP protein.
 XX


```
CC complete plants).
XX
SQ Sequence 2295 BP; 668 A; 447 C; 597 G; 583 T; 0 other;

Query Match      31.9%; Score 151.4; DB 21; Length 2295;
Best Local Similarity 59.0%; Pred. No. 2.2e-40;
Matches 279; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

Qy 1 acaagatgtattatgcagcgtttcaaggttcaacgaatttttggaattcttaagcaa 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1589 aataggatgttttggcagtggtcttcggttgagcaactggctgggaccccttctcag 1648

Qy 61 gggcctaagaatgcacacctgcagcctgttactggtctatatgttcggcaaaagccctc 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1649 ggtctggaatgcgtctctcctgaagcactgttactggtttacatgtttggcaagggtgt 1708

Qy 121 tactttgcagatctagtaagaagagcgcaatactactgtttatgttgatagaataatcct 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1709 tactttgtgacatgttttcaagagtgcaaacattattgtctacgctctggaagcatgaga 1768

Qy 181 gtagtttgatgctctttctgaggttgccttttaggagacatgtatgaactaaagaagcc 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1769 tctggagtactgctttatgtgaggttgcattggcattggcgatgaatgactgaatgca 1828

Qy 241 acgtccatg---acaaacctccaagagggaagcattcgcccaagggtattagcgaaccc 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1829 gattacatgctataacacctgcccaagggaataataagatccaaggaggtgtgcaaca 1888

Qy 298 gtgcactgagtcagcaggtttgtgaagtggagggatgatgtcgtagtctccctcgccaag 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1889 gcacctaaacatggtcgagtcgaagtcgctgacgagtgtgtgttctcccttggcgaa 1948

Qy 358 ccggtgcacatcataatagagctctgaactcatgtacaaatgagatcatcgtctacaac 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1949 cccaacaggaaccttccaaagggtgtgtgtgttataatgagtagatcagtgtagcaac 2008

Qy 418 acatccaggtggaagtcagcttctgtggaagtggtgttccatccacaagag 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2009 gtagaccagataagaalgcggtgtgtcttacctatgtttaacttcaatttcaagag 2061

RESULT 6
AAZ23799
ID AAZ23799 standard; cDNA; 3580 BP.
XX
AC AAZ23799;
XX
DT 18-JAN-2000 (first entry)
XX
DE Human poly(ADP-ribose)polymerase cDNA.
XX
KW Gene therapy; poly(adenosine diphosphate-ribose) polymerase; treatment;
KW PARP; antitumor; nuclear DNA repair; proliferating cell; DNA damage;
KW protection; genomic instability; cancer; prevention; human; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 140..3184
FT /*tag= a
FT /product= "poly(ADP-ribose) polymerase"
PN DE19080889-A1.
XX
PD 09-SEP-1999.
XX
PF 03-MAR-1998; 98DE-1008889.
XX
PR 03-MAR-1998; 98DE-1008889.
XX
FA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
```

Buerkle A, Meyer R;
WPI; 1999-509563/43.
P-PSDB; AAY33699.
New gene therapy vector expressing poly(adenosine diphosphate-ribose)-polymerase for treating or preventing tumors -
Claim 4; Fig 1; 12pp; German.
This invention describes a novel vector (A), suitable for gene therapy, which contains a DNA insert (I) that encodes a practically complete poly(adenosine diphosphate-ribose) polymerase (PARP) which has antitumor activity. PARP is a nuclear DNA repair enzyme (activated by strand breakage), that allows recovery of proliferating cells from the toxic effects of DNA damage (caused by alkylating or oxidizing agents or radiation), and protects cells against such damage or other causes of genomic instability. (A) are used for treatment of cancer and for cancer prevention in subjects at high risk (e.g. those with tumor-associated genetic defects). Tumor cells treated with (A) show an increased tendency to die when treated with radiation or chemotherapeutic agent, and in cells that survive this treatment they inhibit genomic instability, so should reduce the likelihood of further development of both malignant cells and resistance to chemotherapy. This sequence encodes the human poly(ADP-ribose) polymerase described in the invention.
Sequence 3580 BP; 986 A; 823 C; 996 G; 775 T; 0 other;

Query Match 30.1%; Score 142.6; DB 20; Length 3580;
Best Local Similarity 58.0%; Pred. No. 2.4e-37;
Matches 272; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

Qy 1 acaagatgtctattatgcagcgtttcaaggttgacgaatttttggaattcttaagcaa 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2705 aaccgaagattgctgtggcacgggtgccaggaccacaaactttgtggatctgtcccaag 2764

Qy 61 gggcctaagaatgcacacctgcagcctgttactggtctatatgttcggcaaaagccctc 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2765 ggtcttcgtagcctcctgaagcgccctgacaggctacatgtttgtgaaggatc 2824

Qy 121 tactttgcagatctagtaagaagagcgcaatactgttattgtgtaggaaataatcct 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2825 tattcgtgacatggtctcccaaggagtgccactactgcacactcagggagaccca 2884

Qy 181 gtaggtttgatgcttcttctgaggttcttaggagacatgtatgaactaaagaagcc 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2885 ataggcttaactcctgttgggaagattgaccttgaaacatgtatgaactgaagcagct 2944

Qy 241 acgtccatggacaacacctccaagagggaagcattcgcccaagggtattagcgaacccgtg 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2945 tcacatcagcaagttaccgaaggcgcaagcagtgccaagggtttgggcaaaactacc 3004

Qy 301 ccactggatcagagttgtgaagtggaggtatgtatgctgtccctgcggcgaagccg 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3005 cctgacctctcag---ctaacattatgctggatggtgtagacgttctctcttggaccggg 3061

Qy 361 gtgccatcatcaattagggagctctgaactcatgacaatgagtagcatcgtctacacaca 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3062 atttcattggtgtgaatgacacctctctactatataacagtagcatattgtatgatt 3121

Qy 421 tccagggtgaagatgcagttcttctgctgaaggtgcgttttccatcacaaga 469
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3122 gctcaggtaaatcgtgaagtatctgctgaactgaaatcgaatcgaatttaaga 3170

RESULT 7
AAZ13732
ID AAT13732 standard; DNA; 3793 BP.
XX
AC AAT13732;
XX
DT 31-OCT-1996 (first entry)

XX DNA-binding domain of poly(ADP-ribose) polymerase coding sequence.
 DE PARG; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;
 KW tumour treatment; DNA repair; over-expression; ss.
 XX Homo sapiens.
 XX
 FH Location/Qualifiers
 FT CDS
 FT
 FT /*tag= a
 FT /product= poly(ADP-ribose)_polymerase
 FT /note= "PARG"
 FT 67..1220
 FT /*tag= b
 FT /note= "encodes DNA-binding domain"
 XX
 PN W09618737-A2.
 XX
 PD 20-JUN-1996.
 XX
 XX 15-DEC-1995; 95WO-DE01817.
 XX
 PR 16-DEC-1994; 94DE-444949.
 XX
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 XX Buerkle A, Kuepper J, Zur Hausen H;
 XX
 XX WPI; 1996-300654/30.
 DR P-PSDB; AAR99642.
 XX
 XX Vectors contg. insert encoding DNA-binding domain of
 PT poly(ADP-ribose) polymerase - useful for gene therapy, esp. of
 PT tumours
 XX
 XX Claim 4; Fig 1; 22pp; German.
 XX
 CC The present sequence encodes a poly(ADP-ribose) polymerase (PARG) contg.
 CC a DNA-binding domain (DBD). The DNA fragment from -29 to +1127
 CC (nucleotides 67-1220 of this sequence) encoding the DBD can be inserted
 CC into vectors which are used for gene therapy. Over-expression of the DBD
 CC inhibits the DNA repair function of PARG, so the vectors are useful for
 CC gene therapy or tumours, esp. in combination with conventional chemo-
 CC and/or radiotherapy.
 XX
 SQ Sequence 3793 BP; 1049 A; 847 C; 1034 G; 863 T; 0 other;

Query Match 30.1%; Score 142.6; DB 17; Length 3793;
 Best Local Similarity 58.0%; Pred. No. 2.5e-37;
 Matches 272; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 1 acaagatgctattatggcagcgttcaaggttgacgaattttgtgggaattcttagtcaa 60
 DB aaccgaagattgctgtgacggttcacgagaccacaaactttgctggatctgtcccg 2717
 QY 61 gggcctaagaattgacacctctgagccactgttactggtatattgttcgcaagcctc 120
 DB 718 ggtcttcgtagcccgctgaagcgcctgtgacagggctacatgtttgtgaaaggatc 2777
 QY 121 tactttcagatctatgaagcagagcgacaaatactgttattgtgtaggaataatcct 180
 DB 2778 tatttcgtagcatggtctccaagatgccaactactgccaatactgctcaggagacca 2837
 QY 181 gtagggttgatgctctttctgaggttgcttttagggagacatgtatgaactaaagaagcc 240
 DB 2838 ataggcttaactcgtgtgggagaagttgccttggaaactgtatgaactgaagcagct 2897
 QY 241 acgtccatggacaaactccaagagggaagcattcgacaaaggattaggcaaacccgtg 300
 DB 2898 tcacatacagcaagttaaccacaaaggcgaagcagctgtcacaagggttcgggcaaacctacc 2957

QY 301 ccactggagtcagagtttgtgaagtggagggatgatgctagttccctgcggaagcgcg 360
 DB 2958 cgtgaccttcag---ctaacattagtctggtggtgtagacgttccctcttgggaccggg 3014
 QY 361 gtccatcatcaattaggagcctctgaactcatgtacaaatgagtagcatcgtctacaacaca 420
 DB 3015 attcatctggtgtgaatgacacctctctactataacagtagtacattgtctatgatatt 3074
 QY 421 tccagggtgaagatgcagttcttctgctgaaggtgcgtttccatcaacaaga 469
 DB 3075 gctcaggtaaatctgaagtattctgtgaaactgaaattcaattttaaga 3123

RESULT 8
 AAF63954
 ID AAF63954 standard; DNA; 3045 BP.
 XX
 AC AAF63954;
 XX
 DT 05-APR-2001 (first entry)
 XX
 DE Human tankyrase2 related coding sequence SEQ ID NO: 136.
 XX
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
 KW inflammatory disorder; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200100849-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000WO-US17827.
 XX
 PR 29-JUN-1999; 99US-0141582.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 XX
 DR WPI; 2001-102896/11.
 DR P-PSDB; AAB66296.
 XX
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
 XX inflammatory and autoimmune disorders -
 PS Example 2; Page 203-207; 242pp; English.
 XX
 CC The present invention provides the protein and coding sequence for the
 CC human tankyrase2 protein. This is found in two different versions,
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
 CC polyADP-ribosylation activity and is involved in the modification of
 CC TRF1, which is a telomere-specific binding protein. The regulation of
 CC telomere length, in which TRF1 has a role, is linked to ageing and
 CC cancer. The sequences are useful in the treatment of cancers and
 CC inflammatory disorders.
 XX
 SQ Sequence 3045 BP; 863 A; 710 C; 851 G; 621 T; 0 other;

Query Match 29.4%; Score 139.4; DB 22; Length 3045;
 Best Local Similarity 57.6%; Pred. No. 2.7e-36;
 Matches 270; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

QY 1 aacaagatgctattatggcagcgttcaaggttgacgaattttgtgggaattcttagtcaa 60
 DB 2566 aaccgaagattgctgtgacggttcacgagaccacaaactttgtggatctgtcccg 2625
 QY 61 gggcctaagaattgacacctctcctcaggacacctgttactgctatattgttcgcaagcctc 120
 DB 2626 ggtcttcgtagcccgctgaagcgcctgtgacaggtcattgtttgtgaaaggatc 2685

QY 121 taatttcagatctagtaagcagagcgcaataactgttatgtgtaggaataatct 180
 Db 2686 tatttcgctgacatggtctcccaagagtgccaaactactaccacacgtctcagagagacc 2745
 QY 181 taggtttgagctctcttcttctgaggtgttttaggagacatgtatgaactaaagaagcc 240
 Db 2746 atagggttaactcctgttgaggagaagtgcccttgaaacatgtatgaactgaagcacgt 2805
 QY 241 acgtccatggacaacctccaaagaggaagcatttggaccacaaaggtattagcacaacccgtg 300
 Db 2806 tcacatcatcagcaggttaaccagggcaagcacagtgtaaaaggtttggcacaactacc 2865
 QY 301 ccactggagtcagagttttagtaagtgaggagtgatgtctgttagtccctgcgcgaagccg 360
 Db 2866 cctgacccctcag---ctaacattagctggatggctgagacgttccctcttgggaccggg 2922
 QY 361 gtccatcatcaattagggagctctgaactcatgtacaatgagtagacatcgtctacaacaca 420
 Db 2923 atttcactcgtgtgatagacacctctctactataataacgagtagacattgtctatgatatt 2982
 QY 421 tccagggtgaagatgcagttcttctgtcgaaggtgcgtttccatcacacaaga 469
 Db 2983 gctcaggtaaatcgaagtatcgtgaaactgaaactgaaattcaattttaaga 3031

RESULT 9
 AAT04221
 ID AAT04221 standard; DNA; 5345 BP.
 XX
 AC AAT04221;
 XX
 DT 31-MAR-1996 (first entry)
 XX
 DE 5.3 kb EcoRI fragment containing poly(ADP-ribose)-polymerase gene.

XX
 KW DNA primer; PCR; polymerase chain reaction;
 KW poly(ADP-ribose)-polymerase; cancer diagnosis; chromosome-13;
 KW pseudogene; Burkitt's lymphoma; B-follicular cell lymphoma;
 KW small cell lung carcinoma; colorectal carcinoma; tumor;
 KW breast carcinoma.
 XX
 OS Homo sapiens.

XX
 PN US5449605-A.
 XX
 PD 12-SEP-1995.
 XX
 PF 14-OCT-1988; 88US-0257696.
 XX
 PR 06-APR-1993; 93US-0044618.
 PR 14-OCT-1988; 88US-0257696.
 XX
 PA (GEOU) UNIV GEORGETOWN.

XX
 PI Cherney B, Lyn D, Smulson ME;
 XX
 DR WPI; 1995-327692/42.
 XX
 PT Allele-specific primers for detecting pre-disposition to cancer
 PT caused by a deletion in chromosome 13 pseudo-gene for poly
 PT (ADP-ribose) polymerase
 XX
 PS Disclosure; Page 41-46; 47pp; English.

XX
 CC This sequence corresponds to a 5.3 kb EcoRI fragment that contains
 CC the poly(ADP-ribose)-polymerase pseudogene.
 XX
 SQ Sequence 5345 BP; 1598 A; 1142 C; 1396 G; 1209 T; 0 other;

Query Match 29.1%; Score 137.8; DB 16; Length 5345;
 Best Local Similarity 57.4%; Pred. No. 1.2e-35;
 Matches 269; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 1 aacaagatgtattatgacacggttcaaggttgacgaatttttggaattcttagtcaa 60
 Db 3519 aacigaaggttgcgtggtgcatgggtccagaccacaaactttgtggtatcttccctg 3578
 QY 61 gggctaaagaattcacctcctcagagcaccctgttactggctatatgttcgcgaagccctc 120
 Db 3579 ggtcttggatagccctgcctgaagcacctgtgatggctacatgttttggcaagtgatc 3638
 QY 121 tactttgcagatctagtaagaagagcgcaataactactgttatgtggataggaataatcct 180
 Db 3639 tatttcgctgaccttctgtctccaaagtgccaaagcactccatacatcttaggaagaccca 3698
 QY 181 taggtttgagctcttcttcttctttaggttctttaggagacatgtatgaactaaagaagcc 240
 Db 3699 atagggttaactcctgtcggaaagagttgccccttgaaacgtgtggaactgaagcatgct 3758
 QY 241 acgtccatggacaacctccaaagagggagagcattcgcacaaggttaggcacaacccgtg 300
 Db 3759 tcacatcatcagcaagttaaccgaagggcagcacagtgtaaaaggttttggcacaactact 3818
 QY 301 ccactggagtcagagttttaggaagtgaggtgatgtctgttagtccctgcgcgaagccg 360
 Db 3819 cctgaccttctcag---ctagtatccactcgtggtgtgaggttctccttgggaccagg 3875
 QY 361 gtgcatcatcaattagggagctctgaactcatgtacaatgagtagacatcgtctacaacaca 420
 Db 3876 gtttcactcgtgtgtaagacacctgtctactgtataatgagtagacatgtctatgatatt 3935
 QY 421 tccagggtgaagatgcagttcttctgtcgaaggtgcgtttccatcacacaaga 469
 Db 3936 gctcaggtaaatcgaatatctgtgaaactgaaactgaaattcaattttaaga 3984

RESULT 10
 AAT04219
 ID AAT04219 standard; DNA; 2682 BP.
 XX
 AC AAT04219;
 XX
 DT 31-MAR-1996 (first entry)
 XX
 DE Poly(ADP-ribose)-polymerase alpha-allele DNA from chromosome-13.

XX
 KW DNA primer; PCR; polymerase chain reaction;
 KW poly(ADP-ribose)-polymerase; cancer diagnosis; chromosome-13;
 KW pseudogene; Burkitt's lymphoma; B-follicular cell lymphoma;
 KW small cell lung carcinoma; colorectal carcinoma; tumor;
 KW breast carcinoma; ss.

XX
 OS Homo sapiens.
 XX
 PN US5449605-A.
 XX
 PD 12-SEP-1995.
 XX
 PF 14-OCT-1988; 88US-0257696.
 XX
 PR 06-APR-1993; 93US-0044618.
 PR 14-OCT-1988; 88US-0257696.
 XX
 PA (GEOU) UNIV GEORGETOWN.

XX
 PI Cherney B, Lyn D, Smulson ME;
 XX
 DR WPI; 1995-327692/42.
 XX
 PT Allele-specific primers for detecting pre-disposition to cancer
 PT caused by a deletion in chromosome 13 pseudo-gene for poly
 PT (ADP-ribose) polymerase
 XX
 PS Claim 4; Page 31-36; 47pp; English.

CC This sequence is of the poly(ADP-ribose)-polymerase alpha-allele.
 CC A DNA primer specific to this sequence may be used in a PCR
 CC amplification method to differentiate between the alpha- and beta-
 CC alleles of the processed pseudogene.
 XX
 SQ Sequence 2682 BP; 772 A; 618 C; 711 G; 581 T; 0 other;

Query Match 28.7%; Score 136.2; DB 16; Length 2682;
 Best Local Similarity 57.1%; Pred. No. 3e-35;
 Matches 268; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1 aacaagatgctattatggcagcgttcaaggttgacgaattttgtgggaatttttagtcaa 60
 DB 2050 aactgaaggttgctgtgctggtccaggaccaccactttgctggatcctgctccg 2109
 QY 61 gggctaaagattgacactcctgagcagcctgttactggtctatattgttcggcaagcctc 120
 DB 2110 ggtcttgtagcctgctgagcagcctgtgagggctacatgttggtaagatgatc 2169
 QY 121 tacttgcagatctagtaagcagagcgcacaaactgttattgttgataggaaataatct 180
 DB 2170 tatttcgtgatttctctccagagtgccaaagactgcacacatctttaggaagacc 2229
 QY 181 gtagggttgatcttcttctgaggttgcttttagagacatgtatgaactaaagaaacc 240
 DB 2230 atagggttaactctgtcgaagaagttgcttggaaacgtgtgaaactgaagcatgct 2289
 QY 241 acgtcatgagcaaacctccagaggaagcattgcacaaaggttaggacaaacctg 300
 DB 2290 tcacatcagcaagttaccaaaggcagacagtcagcaggttttggcacaactact 2349
 QY 301 ccactggaatcagagttgtgagtgagggtgagtgctgttccctgcggcaagcgc 360
 DB 2350 cctgacacttctcag---ctagatcccaactggtggttagaggttccctctgggaccag 2406
 QY 361 gtgcatcatcaattaggagctctgaactcagtcacatgtagtgcacatgcttacaacaca 420
 DB 2407 gttcatctggtgtgaatgacactctctactgttataatgtagtactgtctatgatatt 2466
 QY 421 tccagggtgaagatcagcttctgtcgaaggtggtggttcccatcaaga 469
 DB 2467 gtcaggttaaatctgaatatctctgctgaactgaactgaattcaatttaaga 2515

RESULT 11
 AAC85320
 ID AAC85320 standard; cDNA; 864 BP.
 XX
 AC AAC85320;
 XX
 XX
 DT 29-MAR-2001 (first entry)
 XX
 DE 3' end fragment of hparp2, 3'-hparp2.
 XX
 XX Human; poly(ADP-ribose) polymerase; hparp2; oxidative stress; ARDS;
 KW inflammation; ischemic stroke; hemorrhagic shock; myocardial ischemia;
 KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
 KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
 KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
 KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;
 KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
 KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
 KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;
 KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
 KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
 KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;
 KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
 KW graft versus host disease; allograft rejection; cystic fibrosis;
 KW chronic glomerulonephritis; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;
 KW inflammatory dermatitis; contact dermatitis; atopic dermatitis;
 KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;

KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;
 KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;
 KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;
 KW expressed sequence tag; EST; RACE; PCR; amplify; primer;
 KW polymerase chain reaction; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..864
 FT /*tag= a
 FT /product= "hPARP2 C-terminal fragment"
 FT
 FT W0200077179-A2.
 PN 21-DEC-2000.
 PD 16-JUN-2000; 2000WO-US16629.
 XX 16-JUN-1999; 99US-0139543.
 PR (ICOS-) ICOS CORP.
 XX
 PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
 PI WPI: 2001-025335/03.
 XX P-PSDB: AAB47031.
 DR
 DR
 DR
 XX
 XX New human poly(ADP-ribose) polymerase for treating inflammatory,
 PT neurological, cardiovascular, or neoplastic tissue growth disorders,
 PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
 PT metastasis
 XX
 XX Example 2; Page 101; 129pp; English.
 XX
 CC This sequence represents the 3' fragment of the human parp2
 CC (poly(ADP-ribose) polymerase) cDNA. This sequence was amplified
 CC using the primers given in AAC85315-17. DNA derived from a testis
 CC cDNA library was used as a template. The protein of the invention,
 CC hPARP2, causes the covalent addition of polymers of ADP-ribose to
 CC protein targets. hPARP2 activity is induced in many instances of
 CC oxidative stress or during inflammation where there is direct
 CC damage to the DNA. hPARP2 may be used to identify antagonists
 CC which may be used to treat a human having a disorder mediated by
 CC PARP2 activity, such as, inflammatory, neurological, cardiovascular,
 CC or neoplastic tissue growth disorders. hPARP2 and antibodies to it,
 CC can also be used to diagnose these conditions.
 XX
 SQ Sequence 864 BP; 250 A; 207 C; 193 G; 214 T; 0 other;

Query Match 25.9%; Score 122.6; DB 22; Length 864;
 Best Local Similarity 56.9%; Pred. No. 6.2e-31;
 Matches 268; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 1 aacaagatgctattatggcagcgttcaaggttgacgaattttgtgggaatttttagtcaa 60
 DB 376 aacaggatgcttctatggtcaggttccaggatgagtaactgggtgggaattcttgagccat 435
 QY 61 gggctaaagattgacactcctgagcagcctgttactggtctatattgttcggcaagcctc 120
 DB 436 gggctcgaattgcccaccctgaagctcccatcacaggttacctgtttgggaaggaatc 495
 QY 121 tacttgcagatcttagtaagcagcagcagcacaactcttatgttgataggaaataatct 180
 DB 496 tacttgcagatcttcttccaaagagtgcccaactctgttctcctcgcctctaaagaa 555
 QY 181 gtagggttgatgcttcttctgtagggttctttaggagacatgtatgaactaaagaaacc 240
 DB 556 acaggactgctgctcttctatcagaggtagctcttagtgcagtgtaataactactagagcc 615
 QY 241 acgtcctcaggacaa---acctccaagaggggaagcattcgaccaggtattaggcaaaccc 297

Db 616 aatcctaagccgaaggattgtctcaaggtaaacacatagcaccacgaagggtggcgaagatg 675
 Qy 298 gtccactgagtcagagtttg-----tgaagtgaaggatgatgtcgtagttccctgc 351
 Db 676 gtccactgagtcagagtttg-----tgaagtgaaggatgatgtcgtagttccctgc 735
 Qy 352 ggaagcgggtgcatcatcaatagagctctgaactcatgtacacatgacatgacatcgtc 411
 Db 736 agtgacacagaattctgaatccagatggttatccctcaactacaatgaatatattga 795
 Qy 412 tacaacacatccacagtggaagatgcagttctgtcgaagggtgcgtttccat 462
 Db 796 tataaccccacacaggtccgtatgcgtacaccttttaagagttcagttta 846

RESULT 12

AAF59996
 ID AAF59996 standard; cDNA; 1566 BP.

XX AAF59996;

DT 11-MAY-2001 (first entry)

XX Human poly(ADP-ribose) synthetase shbPARS2 cDNA.

XX Human; poly(ADP-ribose) synthetase; shbPARS2; vaccine; drug screening;
 KW ischaemic disorder; cerebral ischaemia; cardiac ischaemia;
 KW myocardial infarction; stroke; inflammation; autoimmune disease;
 KW diabetes; multiple sclerosis; neurodegenerative disease;
 KW Parkinson's disease; Alzheimer's disease; chromosome localisation; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..1566

FT /*tag= a
 FT /product= "Human poly(ADP-ribose) synthetase shbPARS2"

PN WO200112645-A1.

XX 22-FEB-2001.

XX 10-AUG-2000; 2000WO-US21775.

XX 12-AUG-1999; 99US-0373441.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;

XX WPI; 2001-211196/21.

XX P-PSDB; AAB60693.

XX New human polypeptide of the polyADPribose synthetase family for use in
 PT diagnostic assays and for screening modulators used for preventing and
 PT treating inflammation, autoimmune disease and Alzheimers disease

XX Claim 2; Page 21; 30pp; English.

XX This sequence represents cDNA encoding a novel human poly(ADP-ribose)
 CC synthetase, shbPARS2. The invention also relates to fragments,
 CC variants and sequences with at least 95% identity to the shbPARS2
 CC protein or nucleotide sequence; expression systems and host cells
 CC comprising an shbPARS2 nucleic acid sequence; The recombinant expression
 CC of shbPARS2; and an antibody specific for shbPARS2. shbPARS2 proteins
 CC and nucleotides are useful as vaccines for inducing an immunological
 CC response in a mammal. The shbPARS2 protein is useful for identifying
 CC compounds which inhibit or stimulate its activity or expression level.
 CC Such agonists and antagonists of shbPARS2 are useful for treating human
 CC diseases including ischaemia and ischaemic tissue injury (e.g., cerebral
 CC and cardiac ischaemia, myocardial infarction, stroke), inflammation,
 CC autoimmune disease (e.g., diabetes, multiple sclerosis) and
 CC neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's

CC disease). shbPARS2 nucleic acids are useful as diagnostic reagents for
 CC detecting mutations in the associated gene; as hybridisation probes; and
 CC to isolate full-length shbPARS2 cDNAs and shbPARS2 genomic clones; and
 CC as an immunogen to produce antibodies for therapeutic use. shbPARS2
 CC proteins, nucleotides and antibodies are also useful in screening methods
 CC for detecting the effect of added compounds on the production of mRNA and
 CC protein in cells.

XX Sequence 1566 BP; 490 A; 328 C; 381 G; 367 T; 0 other;

Query Match 25.9%; Score 122.6; DB 22; Length 1566;
 Best Local Similarity 56.9%; Pred. No. 8.3e-31;
 Matches 268; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

Qy 1 aacaagatgctattatggcacggtttcaaggttgacgaattttgtgggaattcttagtcaa 60
 Db 1078 aacagatgcttctatgcatgtttccagatgagtaactgggtgggaattcttagtcaa 1137
 Qy 61 gggctaagaattgcaoctctctgagcaccctgttactggctatatgttcggcaaaagcctc 120
 Db 1138 gggcttgaattgcccacactgaagctccatcacaggttacctgttgggaaggaatc 1197
 Qy 121 tactttgcagatctagtaagcaagagcgacacatactgttatgtgtaggaataactct 180
 Db 1198 tactttgctgacatctctccaaagagtgccaattactgcttgcctcctcctaaagaat 1257
 Qy 181 gtaggtttgatgcttcttctgaggttgccttaggagacatgtatgaactaaagaagcc 240
 Db 1258 acaggactgctcttctatcacaggttagctctgagtcagtgtaagtaactactagggcc 1317
 Qy 241 acgtccatgacaa----acctccaaagagggaagcattccacaaagagattagcaaaacc 297
 Db 1318 aatccaaaggcgaaggattgttcaaggtaaacatagcaccaaagggtggcgaagt 1377
 Qy 298 gtccactgagtcagagtttg-----tgaagtgaaggatgatgtcgtagttccctgc 351
 Db 1378 gtccactgagtcagagtttg-----tgaagtgaaggatgatgtcgtagttccctgc 1437
 Qy 352 ggaagcgggtgcatcatcaatagagctctgaaactcatgtacacatgacatcgtc 411
 Db 1438 agtgacacagaattctgaatccagatggttatccctcaactacaatgaatatattga 1497
 Qy 412 tacaacacatccacaggtgaagatgcagttctgtcgaagggtgcgtttccat 462
 Db 1498 tataaccccacacaggtccgtatgcgtacaccttttaagagttcagttta 1548

RESULT 13

AAC85303

ID AAC85303 standard; cDNA; 1814 BP.

XX AAC85303;

XX 29-MAR-2001 (first entry)

XX hparp2 cDNA.

XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;
 KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;
 KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;
 KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;
 KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;
 KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;
 KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;
 KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;
 KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;
 KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;
 KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;
 KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;
 KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;
 KW graft versus host disease; allograft rejection; cystic fibrosis;

chronic glomerulonephritis; inflammatory bowel disease; Crohn's disease; ulcerative colitis; necrotizing enterocolitis; inflammatory dermatosis; contact dermatitis; atopic dermatitis; psoriasis; urticaria; fever; myalgia; meningitis; encephalitis; Sjögren's syndrome; alcoholic hepatitis; bacterial pneumonia; hypovolemic shock; type 1 diabetes mellitus; hypersensitivity; leukocyte dyscrasia; thermal injury; cytokine-induced toxicity; ds.

OS Homo sapiens.

...	key	Location/Qualifiers
FH	CDS	63..1814
FT		/*tag= a
FT		/product= "hPARP2"

WO200077179-A2

21-DEC-2000.

16-JUN-2000; 2000WO-US16629.

16-JUN-1999; 99US-0139543.

PA (ICOS-) ICOS CORP.

PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL:

WPI; 2001-025335/03.

DR P-PSDB; AAB47029.

PT New human poly(ADP-ribose) polymerase for treating inflammatory,
PT neurological, cardiovascular, or neoplastic tissue growth disorders,
PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte
PT metastasis.

PS Claim 1; Page 91-93; 129pp: English.

This sequence encodes human poly(ADP-ribose) polymerase (hPARP2). This protein causes the covalent addition of polymers of ADP-ribose to protein targets. hPARP2 activity is induced in many instances of oxidative stress or during inflammation where there is direct damage to the DNA. hPARP2 may be used to identify antagonists which may be used to treat a human having a disorder mediated by PARP2 activity, such as, inflammatory, neurological, cardiovascular, or neoplastic tissue growth disorders, e.g. ischemic stroke, hemorrhagic shock, myocardial ischemia or infarction, transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic shock, gram negative or positive sepsis, toxic shock syndrome; multiple organ injury syndrome secondary to septicemia, trauma, or hemorrhage; allergic or vernal conjunctivitis, uveitis, thyroid-associated ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis, allergic rhinitis, ARDS, chronic obstructive pulmonary disease, silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis, pneumonia, bronchiectasis, pulmonary oxygen toxicity; reperfusion injury of the myocardium, brain or extremities; cystic fibrosis; keloid formation, scar tissue formation; atherosclerosis; systemic lupus erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's syndrome; graft versus host disease, allograft rejection; chronic glomerulonephritis; inflammatory bowel disease, Crohn's disease, ulcerative colitis, necrotizing enterocolitis; inflammatory dermatoses, contact or atopic dermatitis, psoriasis, urticaria, fever and myalgias due to infection; meningitis, encephalitis, and brain and spinal cord injury due to minor trauma; Sjogren's syndrome; diseases involving leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia; antigen-antibody complex mediated diseases; hypovolemic shock; Type 1 diabetes mellitus; acute and delayed hypersensitivity; disease states due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte transfusion associated syndromes; and cytokine-induced toxicity. hPARP2 and antibodies to it, can also be used to diagnose these conditions.

Sequence 1814 BP; 558 A; 389 C; 462 G; 405 T; 0 other;

Query Match 25.9%; Score 122.6; DB 22; Length 1814;
Best Local Similarity 56.9%; Pred. No. 9e-31;
Matches 268; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

Qy	1	aacaagatgctattatggcacggttcaaggttgacgaatttttgggaattcttagtcaa	60
Db	1326	aacaggatgtcttatggcatggttccaggatgagtaactgggtgggaattctgagccat	1385
Qy	61	ggctaaagaattgcacctcctgaggcacctgttactgctatatgtttgcgcaaggcctc	120
Db	1386	ggcttcgaattgcccaactgaagctccatccacaggtttacatgtttgggaaggaaac	1445
Qy	121	tactttgcagatctatgaagcaagcgcacaaactctgttatgtggatagggaataactct	180
Db	1446	tactttgtgacatgcttcccaagagtgccaattactgtttgctctcgcgtaaagaat	1505
Qy	181	gtaggtttgatgctcttcttctgaggttgcctttaggagacatgatatgaactaaagaagcc	240
Db	1506	acaggactgctgctcttatacagagtagctctagttagtgaatgaactactagagcc	1565
Qy	241	acgtccatggacaa---acctccaagaggaagcatctgcaccaaggataggcacaacc	297
Db	1566	aatcctlaaggcgaaggattgtctcaagtgaataacatagaccaaggggctggccaagatg	1625
Qy	298	gtgccactggagtcgaaggtttg-----tgaagtggagggatgatgtcgtagttccctgc	351
Db	1626	gtctccagttctgcccacttctcaccttgatggagtagcagtgccattaggaccagca	1685
Qy	352	ggcaagccggtgccatcatcaattaggagcgtcaactcatctgatacgaatgatacatcgtc	411
Db	1686	agtgacacagggaattctgaatccagatggttatcccttcaactacaatgaatatattga	1745
Qy	412	tacaacacatccaggtgaagatgcagttcttctgctgaaggtgcggtttcccat	462
Db	1746	tataaccccaacacaggtccgtatgcgttaccttttaagcttcagtttaagt	1796

RESULT 14

AAZ44287
ID AAZ44287 standard; cDNA; 1843 BP.

AA
AC
AAZ44287:AA
DT 31-MAR-2000 (first entry)

Human brain PARP2 cDNA.

XX
KW PARP: poly(ADP-ribose)

KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW ischemic tissue damage. cc

XX
XX

US Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	3 1715

```

FT      /*tag= a
FT      /product=

```

XX
PN
W09964572-A2XX
PD
16-DEC-1999XX
DE
04 - TRIM - 1000 -

XX
DD
DD
XX

PR 01-MAR-1999; 99DE-1008837.

PA (BADI) BASF AG.

PI Kock M, Hoeger

DR WPI: 2000-087218/07.
 DR P-PSDB; AAF51174.
 XX Novel genes and proteins, antibodies and binding partners useful in
 PT diagnosis and therapy of energy deficiency associated disease
 PT conditions -
 XX Claim 7a; Page 49-52; 96pp; German.
 XX This invention describes novel human and murine poly(ADP-ribose)
 CC polymerase (PARP) homologues, which are characterised by an amino acid
 CC sequence with a functional NAD⁺-binding site and no zinc finger
 CC sequence motif, of general formula CX₂CX₂MX₂C (I). The nucleic acid
 CC sequences, PARP homologues and antibodies are useful for analytic
 CC detection of PARP homologues and for identifying PARP effectors or
 CC binding partners, as well as for determining their effectiveness.
 CC PARP-binding partners are useful for the diagnosis or therapy of a
 CC disease condition, which is the result of a PARP protein, especially an
 CC energy deficiency, which may comprise tissue damage from cell death
 CC following necrosis or apoptosis. The disease condition may be caused
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,
 CC in particular neurotoxic disturbances, etc. This sequence encodes the
 CC human PARP2 protein used in the method of the invention.
 XX
 XX Sequence 1843 BP; 599 A; 377 C; 447 G; 420 T; 0 other;

Query Match 25.9%; Score 122.6; DB 21; Length 1843;
 Best Local Similarity 56.9%; Pred. No. 9e-31;
 Matches 268; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 1 acaagatgctattatgacggttcaaggttgacgaattttgtgggaattcttagtcaa 60
 DB 1227 aacaggatgctctatgctggttccaggatgagtaactgggtgggaattctgagccat 1286
 QY 61 gggctgaagattgacacctctgagggaccctgttactggttatgttcggcgaagggcctc 120
 DB 1287 gggctgaagattgacacctctgagggaccctgttactggttatgttcggcgaagggcctc 1346
 QY 121 tactttcagatctagtaagaagagcgcaataactgtttatgttgataggaataatcct 180
 DB 1347 tactttcagatctagtaagaagagcgcaataactgtttatgttgataggaataatcct 1406
 QY 181 gtaggttgatgctctcttctgaggttgccttttagagacatgtatgaactaaagaagcc 240
 DB 1407 acaggactgctctctatcagaagtgccttagtgcagtgtaactactagagggcc 1466
 QY 241 acgtccatggacaa---acctccaagagggaagcattcgacccaaggattagcgaacacc 297
 DB 1467 aatcctaaggccgaaggattgcttcaagggttaaacatagcacaagggtggcgaagtg 1526
 QY 298 gtgccactggagtcagagatttg-----tgaagtggagggtatgctgtagttccctgc 351
 DB 1527 gctccagttctgcccactctgcacccctgaatggggagacagtgccacttaggacacga 1586
 QY 352 ggcagccggtgcccatactatcattagggagctctgaactcattgatacattgagatcgtc 411
 DB 1587 agtgacacaggaattctgaatccagatggtttataccctcaactcaaatgaatatatgta 1646
 QY 412 tacaacacatcccggtggaagatcaggtcttctgtggaagggtgctttccat 462
 DB 1647 tataaccccaaccagggtcgtgctgacgttacccttttaaaagggttcagttta 1697

RESULT 15

AAC82090
 ID AAC82090 standard; cDNA; 1843 BP.

XX AAC82090;
 AC AAC82090;

DT 02-MAR-2001 (first entry)
 XX Human brain poly-ADP-ribose-polymerase cDNA.

XX Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
 KW cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
 KW anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
 KW antinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
 KW 2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
 KW neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
 KW parkinson's disease; ischemic damage; microinfarction; sepsis;
 KW diabetes mellitus; ss.
 XX Homo sapiens.
 OS
 XX DE19921567-A1.
 PN
 XX 16-NOV-2000.
 PD
 XX 11-MAY-1999; 99DE-1021567.
 PF
 XX 11-MAY-1999; 99DE-1021567.
 PR
 XX (BADI) BASF AG.
 PA
 XX Lubisch W, Sadowski J, Kock M, Hoeger T;
 PI
 XX WPI: 2001-032983/05.
 DR P-PSDB; AAB11480.
 DR
 XX Drugs for inhibiting PARP or especially homologous enzymes comprising
 PT 4-substituted phthalazinone derivatives, useful e.g. for treating
 PT neurodegenerative disease, ischemic damage, tumors or diabetes -
 XX
 PS Example A; Page 9-12; 14pp; German.
 CC This invention describes novel 4-substituted 2H-phthalazin-1-one
 CC derivatives (I) which are used for the treatment or prophylaxis of
 CC diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
 CC EC 2.4.2.30) activity. The products of the invention have nootropic,
 CC neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,
 CC cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,
 CC immunosuppressive, antinflammatory, antirheumatic, antiarthritic,
 CC antidiabetic. (I) are especially used for treating or preventing
 CC neurodegenerative disease or neuronal damage (especially associated
 CC with ischemia, trauma or massive bleeding, especially apoplexy or
 CC spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or
 CC Parkinson's disease), treating or preventing ischemic damage
 CC (specifically renal damage after renal ischemia or during and after
 CC kidney transplantation or heart damage after cardiac ischemia), treating
 CC epilepsy, specifically generalized epileptic attacks (e.g. temporal lobe
 CC tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lobe
 CC and complex partial attacks), treating microinfarction (e.g. during and
 CC after heart valve replacement, aneurysm resectioning and heart
 CC transplantation), revascularization of critically constricted coronary
 CC arteries (e.g. after PCA or by-pass operations) or peripheral arteries
 CC (e.g. leg arteries), treating acute myocardial ischemia and damage during
 CC or after its mechanical or drug-induced lysis and treating tumors and
 CC their metastasis, sepsis and septic shock, inflammatory and rheumatic
 CC disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)
 CC inhibit PARP (i.e. the known form designated PARP1), they especially
 CC selectively and strongly inhibit PARP homologs, specifically the homolog
 CC PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.
 CC with K_i values of 1-20 nM) and high selectivity for PARP2 relative to
 CC PARP1 (generally by a factor of more than 5).
 XX
 XX Sequence 1843 BP; 599 A; 377 C; 447 G; 420 T; 0 other;

Query Match 25.9%; Score 122.6; DB 22; Length 1843;
 Best Local Similarity 56.9%; Pred. No. 9e-31;
 Matches 268; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 1 acaagatgctattatgacggttcaaggttgacgaattttgtgggaattcttagtcaa 60
 DB 1227 aacaggatgctctatgctggttccaggatgagtaactgggtgggaattctgagccat 1286

QY 61 gggctaagaattgcacctctgaggcacctgttactggctatatgttctggcacaaggccctc 120
Db 1287 gggcttcgaattgccccacactgaagctccatcacaggttacatgttgggaaggaaac 1346
QY 121 taactttgcagatctagtaagcaagagcgacacaatactgtttatgtggatagggaataatcct 180
Db 1347 taactttgctgacatgtcttccaaagtgccaattactgtttgctctcgcctaaagaat 1406
QY 181 gtagggttgatgcttcttctgaggttcttttagagacatgtatgaactaaagaagccc 240
Db 1407 acaggactgctctcttaacagaggtagctcttaggtcagtgtaatgaactactagagcc 1466
QY 241 acgtccatggacaa----acctccaagagggaagcattcgaccaagggttagggcaaaacc 297
Db 1467 aatcctaaggccgaaggttcttcaaggtaaacatagcaccacaagggtggcgaagatg 1526
QY 298 gtgccaactggagtcagaggttg-----tgaagtgaggaggtgatgtcgtagttccctgc 351
Db 1527 gctcccagttctgcccactctgacccctgaatggaggtacagtgccattaggaccagca 1586
QY 352 ggcgaagccggtgccatcatcaattaggagctctgaactcactgtacaatgagtacatcgtc 411
Db 1587 agtgacacaggaattctgaatccagatggttataccctcaactacaatgaatatattgta 1646
QY 412 tacaacacatcccaggtgaagatgcagttcttctggaaggtgcgtttccat 462
Db 1647 tataaccccaaccagggtccgtatgcgtacotcttttaaagggttcagtttaat 1697

Search completed: March 7, 2002, 18:16:38
Job time: 13288 sec

Sequence 2, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 58, Appli
Sequence 58, Appli
Sequence 65, Appli
Sequence 65, Appli
Sequence 62, Appli
Sequence 62, Appli
Sequence 63, Appli
Sequence 63, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 64, Appli
Sequence 64, Appli

Sequence 2, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 58, Appli
Sequence 58, Appli
Sequence 65, Appli
Sequence 65, Appli
Sequence 62, Appli
Sequence 62, Appli
Sequence 63, Appli
Sequence 63, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 64, Appli
Sequence 64, Appli

ALIGNMENTS

RESULT 1
US-08-044-618-5
; Sequence 5, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3747 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-044-618-5

Query Match 29.4%; Score 139.4; DB 1; Length 3747;
Best Local Similarity 57.6%; Pred. No. 2.6e-37;
Matches 270; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	139.4	29.4	3747	1	US-08-044-618-5
2	137.8	29.1	5345	1	US-08-044-618-7
3	136.2	28.7	2682	1	US-08-044-618-3
4	114.6	24.2	595	4	US-09-328-111-216
5	108.8	23.0	1592	1	US-08-044-618-4
6	31	6.5	7218	1	US-08-232-463-14
7	30.6	6.5	2757	2	US-08-627-254C-26
8	30.4	6.4	2625	2	US-08-357-533A-1
9	30.4	6.4	2625	3	US-08-459-009-1
10	30.4	6.4	2625	3	US-08-459-951-1
11	29.8	6.3	1160	4	US-07-780-973-15
12	29.4	6.2	3580	4	US-09-081-345-1
13	28.4	6.0	1743	2	US-08-841-178-20
14	28.4	6.0	2947	1	US-08-457-176-1
15	28.4	6.0	2976	1	US-08-457-175-1
16	28.4	6.0	2976	4	US-09-352-159-26
17	28.4	6.0	2976	4	US-09-352-168-26
18	27.6	5.8	328	4	US-09-385-982-227
19	27.6	5.8	758	4	US-09-328-111-685
20	27.2	5.7	3236	4	US-08-927-219-7
21	27.2	5.7	3238	4	US-08-927-219-1
22	27.2	5.7	3238	4	US-08-927-219-3
23	27.2	5.7	3239	4	US-08-927-219-5
24	27.2	5.7	6254	4	US-08-927-219-126
25	27	5.7	323	6	5252465-10
26	27	5.7	1133	1	US-07-750-080A-2
27	27	5.7	1133	1	US-07-750-080A-3

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Searched: 351203 seqs, 113238999 residues 702406
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 16:49:07 ; Search time 122.78 Seconds
(without alignments)
874.333 Million cell updates/sec

Title: US-09-236-995D-3
Perfect score: 474
Sequence: 1 aacaagatgtattatggca.....gtttccatcacaaagagtag 474

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
QY 1 aacaagatgctattatgacagcgttcaaggttgacgaattttgtgggaattcttagtcaa 60
Db 2756 AACCGAAGATTGCTGTGGCAGCGGTCCAGGACCAACCACTTGTGGGATCCTGTCCAG 2815
QY 61 gggtcagaattgacacctcctgagcagcctgttactggtctatgttgcgcaaggcctc 120
Db 2816 GGTCTTCGATAGCCCGCTGAAAGCCCGCTGACAGCGCTACATGTTGTTGTAAGGATC 2875
QY 121 tactttgcagatctagtaagaagcgcacaaatactgttattgtggaataatcct 180
Db 2876 TATTTCGCTGACATGGTCTCCAGAGTGCCCACTACTACCATACGTCCTCAGGAGACCCA 2935
QY 181 gtggtttgagctctcttctgaggttctttaggagacatgtatgaactaaagaagcc 240
Db 2936 ATAGGCTTAATCTCTGTTGGGAGAAGTTGCCCTTGGAAACATGATGAACAGACGCT 2995
QY 241 acgtccatgacaaacctcgaagaggaagcattcgacaaaggattaggcaaaaccgtg 300
Db 2996 TCACATATCAGCAGTTACCCAGGCGCAAGCACAGTGTCAAAGGTTTGGGCAAACTACC 3055
QY 301 ccactggagtcagagttgttgaagtgaggagtgatgctgtagtccctgcgcaagccg 360
Db 3056 CCTGATCCTTCAG--CTAACATTAGTCTGGATGTTGAGACGTTCCCTTTGGACCGGG 3112
QY 361 gtccatcatcaattgagcgtctgaactcatgtacatgtagtacctgtcacacaca 420
Db 3113 ATTTCATCTGGTGTGAATGACACCTCTACTATATATAACGAGTACATTTCTATGATAAT 3172
QY 421 tccagggtgaagtcagctctctgcaaggtgcgttccatcacacaaga 469
Db 3173 GCTCAGGTAAATCTGAAGTATCTGCTGAAACTGAAATTCATTTTAAGA 3221
```

RESULT 2

```
US-08-044-618-7
; Sequence 7, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5345 base pairs
; TYPE: NUCLEIC ACID
```

```
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-044-618-7
```

Query Match

```
Best Local Similarity 29.1%; Score 137.8; DB 1; Length 5345;
Matches 269; Conservative 0; Mismatches 197; Indels 3; Gaps 1;
QY 1 aacaagatgctattatgacagcgttcaaggttgacgaattttgtgggaattcttagtcaa 60
Db 3519 AACTCAAGGTTGCTGTGGCATGGGTCCAGGACCACTTGTGGGATCCTGTCCCTG 3578
QY 61 gggtcagaattgacacctcctgagcagcctgttactggtctatgttgcgcaaggcctc 120
Db 3579 GGTCTTCGATAGCCCGCTGAAAGCCCGCTGACAGCGCTACATGTTGTTGTAAGTATC 3638
QY 121 tactttgcagatctagtaagaagcgcacaaatactgttattgtggaataatcct 180
Db 3639 TATTTCGCTGATCTGTCTCCAAAGAGTGCCCACTACTACCATCTTAGGAAGACCCA 3698
QY 181 gtggtttgagctctcttctgaggttctttaggagacatgtatgaactaaagaagcc 240
Db 3699 ATAGGCTTAATCTCTGCGAAGAGTTGCCCTTGGAAACGTTGTGTAACGACGATGCT 3758
QY 241 acgtccatgacaaacctcgaagaggaagcattcgacaaaggattaggcaaaaccgtg 300
Db 3759 TCACATATCAGCAAGTTACCCAGGCGCAAGCACAGTGTCAAAGGTTTGGGCAAACTACT 3818
QY 301 ccactggagtcagagttgttgaagtgaggagtgatgctgtagtccctgcgcaagccg 360
Db 3819 CCTGATCCTTCAG--CTAGTATCCCACTGGATGTTGAGGTTCTCTTGGGACCAAG 3875
QY 361 gtccatcatcaattgagcgtctgaactcatgtacatgtagtacctgtcacacaca 420
Db 3876 GTTTCATCTGGTGTGAATGACACCTGCTACTGTATATATGAGTACATTTGCTATGATAAT 3935
QY 421 tccagggtgaagtcagctctctgcaaggtgcgttccatcacacaaga 469
Db 3936 GCTCAGGTAAATCTGAAGTATCTGCTGAAACTGAAATTCATTTTAAGA 3984
```

RESULT 3

```
US-08-044-618-3
; Sequence 3, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
```

NAME: FOX, SAMUEL L
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0654.0490001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)466-0800
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2682 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-044-618-3

CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 216
LENGTH: 595
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(595)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-216

Query Match 24.2%; Score 114.6; DB 4; Length 595;
Best Local Similarity 63.3%; Pred. No. 2.6e-29;
Matches 193; Conservative 0; Mismatches 109; Indels 3; Gaps 1;
Qy 1 aacaagatgctattatgacacggttcaaggttgacgaattttgtgggaattcttagtcaa 60
Db 342 AACAGATGCTTCTATGGCATGGTTCAGGATGAGTAACTGCTGGTGGGAATCTTGAGCCAT 283
Qy 61 gggctaaagaattgcacctctgaggaacctgttactggctatatgttcggcaaaaggcctc 120
Db 282 GGGCTTCGAATGTCCTTATGCGCATGGTTCAGGATGAGTAACTGCTGGTGGGAATCTTGAGCCAT 223
Qy 121 tactttgagatctagttaagaacagcgcacaaactactttatgttggaagaaataact 180
Db 222 TACTTTGCTGATCTCTTCCAGAGATGCTTCCAGAGATGCTTCCAGAGATGCTTCCAGAGAT 163
Qy 181 gtaggttgatgctctcttctgaggttcttaggagacatgtatgaactaaagaagcc 240
Db 162 ACAGGACTGCTGCTCTTATCAGAGGTAGTCTAGGTCACTGTAATGAATGAATGAAGCC 103
Qy 241 acgtccatggacaa---acctccaagaggaagcattcgacacaggttaggcaaaagcc 297
Db 102 AATCCTAAGGCCGAGGATTTGCTTCAAGGTAACATAGCACCAAGGGCTGGGCAAGATG 43
Qy 298 gtgcc 302
Db 42 GCTCC 38

Query Match 28.7%; Score 136.2; DB 1; Length 2682;
Best Local Similarity 57.1%; Pred. No. 2.7e-36;
Matches 266; Conservative 0; Mismatches 198; Indels 3; Gaps 1;
Qy 1 aacaagatgctattatgacacggttcaaggttgacgaattttgtgggaattcttagtcaa 60
Db 2050 AACTGAAGGTGCTGTGGCATGGTTCAGGACCAACCACTTCTGCTGGATCTCTGCCCTG 2109
Qy 61 gggctaaagaattgcacctctgaggaacctgttactggctatatgttcggcaaaaggcctc 120
Db 2110 GGTCTTTGGATACCCCTGCTGAAGCACCTGTGATGGCTACATGTTTGGTAAAGTGATC 2169
Qy 121 tactttgagatctagttaagaacagcgcacaaactactttatgttggaagaaataact 180
Db 2170 TATTTGCTGATCTGTCTCAAGAGTGCCACGACTGCTACATCTTAGGAAGACCCA 2229
Qy 181 gtaggttgatgctctcttctgaggttcttaggagacatgtatgaactaaagaagcc 240
Db 2230 ATAGGGTTAATCTGCGGAAGGTGCTTGGAAACGCTGTGAACCTGCTGCTGAGCAAGTCT 2289
Qy 241 acgtccatggacaaacctccaagaggaagcattcgacacaggttaggcaaaagccgtg 300
Db 2290 TCACATATCAGCAAGTATACCCAGGGCAAGCAGACAGTGTCAAGGTTTGGGCAAACTACT 2349
Qy 301 ccactgagtcagatgttgaagtggagagatgctgtagtccctgcggcaagccg 360
Db 2350 CCTGACCTTTCAG---CTAGTATCCCACTGGATGGTGTAGAGGTCTCTTGGGACACAGG 2406
Qy 361 gtgccatcatcaattagaagctctgaactcatgtacaaatgtagtacctacacaca 420
Db 2407 GTTTCATCTGTGTGAATGACACTGCTACTGTATTAATGAGTACATGCTATGATATT 2466
Qy 421 tccaggtgaagatgcagttcttctgctgaaggtgcgttcccatcacaga 469
Db 2467 GCTCAGGTAATCTGAAATATCTCTGAAACTGAAATCAATTTAAGA 2515

RESULT 5
US-08-044-618-4
Sequence 4, Application US/08044618
Patent No. 549605
GENERAL INFORMATION:
APPLICANT: SMULSON, MARK
TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
TITLE OF INVENTION: POLYMERASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Suite 300
CITY: Washington
STATE: D.C.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044,618
FILING DATE: 19930406
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

Query Match 28.7%; Score 136.2; DB 1; Length 2682;
Best Local Similarity 57.1%; Pred. No. 2.7e-36;
Matches 266; Conservative 0; Mismatches 198; Indels 3; Gaps 1;
Qy 1 aacaagatgctattatgacacggttcaaggttgacgaattttgtgggaattcttagtcaa 60
Db 2050 AACTGAAGGTGCTGTGGCATGGTTCAGGACCAACCACTTCTGCTGGATCTCTGCCCTG 2109
Qy 61 gggctaaagaattgcacctctgaggaacctgttactggctatatgttcggcaaaaggcctc 120
Db 2110 GGTCTTTGGATACCCCTGCTGAAGCACCTGTGATGGCTACATGTTTGGTAAAGTGATC 2169
Qy 121 tactttgagatctagttaagaacagcgcacaaactactttatgttggaagaaataact 180
Db 2170 TATTTGCTGATCTGTCTCAAGAGTGCCACGACTGCTACATCTTAGGAAGACCCA 2229
Qy 181 gtaggttgatgctctcttctgaggttcttaggagacatgtatgaactaaagaagcc 240
Db 2230 ATAGGGTTAATCTGCGGAAGGTGCTTGGAAACGCTGTGAACCTGCTGCTGAGCAAGTCT 2289
Qy 241 acgtccatggacaaacctccaagaggaagcattcgacacaggttaggcaaaagccgtg 300
Db 2290 TCACATATCAGCAAGTATACCCAGGGCAAGCAGACAGTGTCAAGGTTTGGGCAAACTACT 2349
Qy 301 ccactgagtcagatgttgaagtggagagatgctgtagtccctgcggcaagccg 360
Db 2350 CCTGACCTTTCAG---CTAGTATCCCACTGGATGGTGTAGAGGTCTCTTGGGACACAGG 2406
Qy 361 gtgccatcatcaattagaagctctgaactcatgtacaaatgtagtacctacacaca 420
Db 2407 GTTTCATCTGTGTGAATGACACTGCTACTGTATTAATGAGTACATGCTATGATATT 2466
Qy 421 tccaggtgaagatgcagttcttctgctgaaggtgcgttcccatcacaga 469
Db 2467 GCTCAGGTAATCTGAAATATCTCTGAAACTGAAATCAATTTAAGA 2515

US-09-328-111-216/c
Sequence 216, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)

APPLICATION NUMBER: US/07/257,696
 FILING DATE: 14-OCT-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: FOX, SAMUEL L
 REGISTRATION NUMBER: 30,353
 REFERENCE/DOCKET NUMBER: 0654.0490001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)466-0800
 TELEFAX: (202)833-8716
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1592 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-08-044-618-4

Query Match 23.0%; Score 108.8; DB 1; Length 1592;
 Best Local Similarity 60.5%; Pred. No. 3.9e-27;
 Matches 179; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
 QY 1 aacaatctattatggcaggttcaaggttgacgaattttgttggaatttttagtcaa 60
 Db 1209 AACTGAAGGTGCTGTGGCATGGTCCAGGACCACCAACTTTGCTGGGATCCTGTCCTG 1268
 QY 61 gggtgaagtgcacacctctgagcagcctgttactggctatatgttcgcaaaagccctc 120
 Db 1269 GGCTTTGGTAGGCTGCTGAGCAGCTGTGATGGGCTAGATGTTGGTAAAGTATC 1328
 QY 121 tacttgcagatctagtgaagcagcgcacaaatactgttatgtgtaggaataatcct 180
 Db 1329 TATTTGCTGTATCTGTCTCCAGAGTGCCAAAGCTGCATACATCTTAGGAAGACCCA 1388
 QY 181 gtagggttgatcttcttctgagttgcttttagagacatgtatgaactaaagaagcc 240
 Db 1389 ATAGGGTTAATCCTGTGGAGAAAGTTGCCCTTGGAAACGGTGTGTGAACCTGAAGCATGCT 1448
 QY 241 acgtcagtgacacacctcccaagagggagcatttcgacaaaggattagcgaacac 296
 Db 1449 TCACATATCAGCAAGTTACCCCAAGGCAAGCACAGTGTCAAGGTTTGGGGCAAAAC 1504

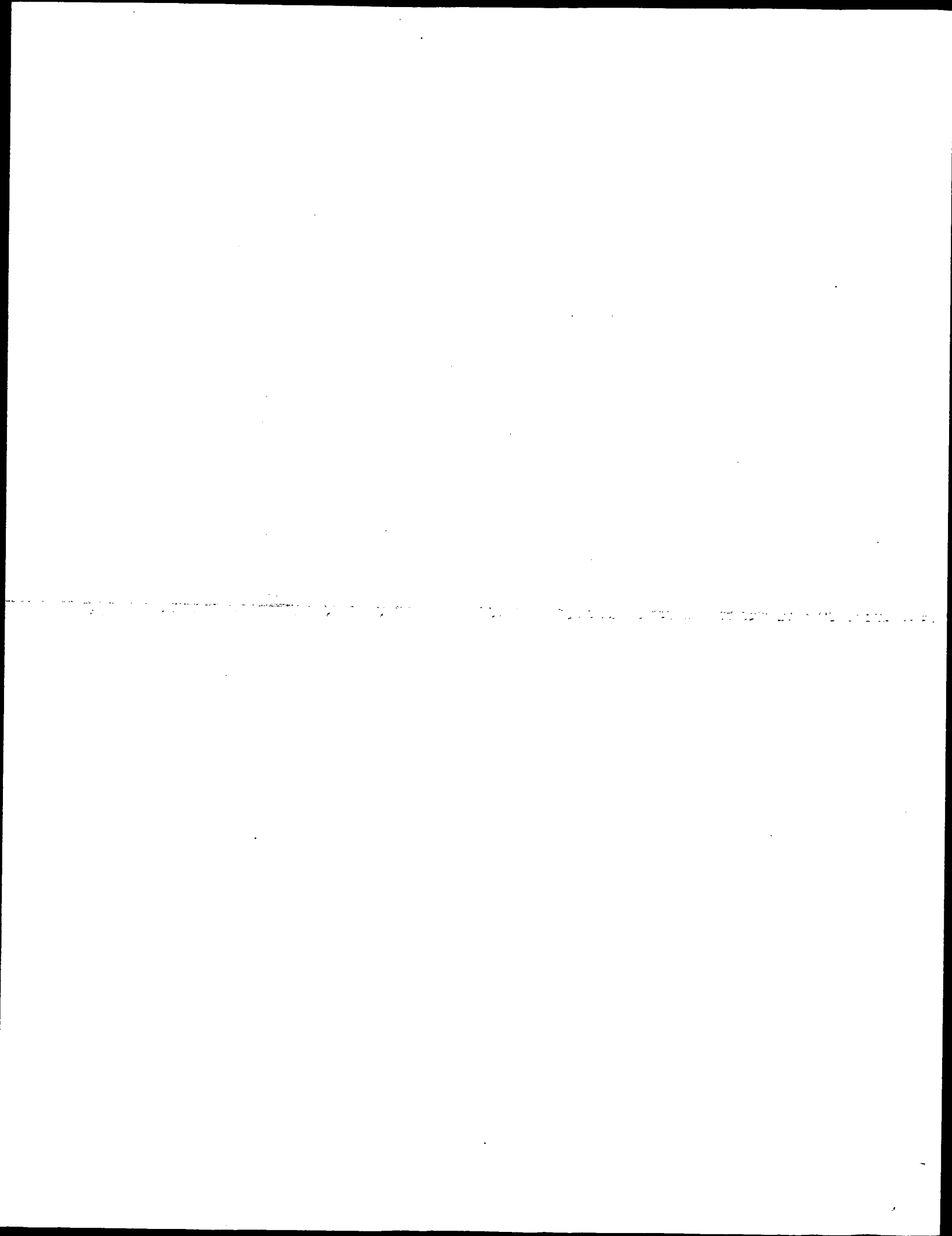
RESULT 6
 US-08-232-463-14/c
 ; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 INMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: pTZgpt-Fls
 US-08-232-463-14

Query Match 6.5%; Score 31; DB 1; Length 7218;
 Best Local Similarity 6.1%; Pred. No. 2;
 Matches 25; Conservative 198; Mismatches 188; Indels 0; Gaps 0;
 QY 28 aggttgacgaattttgtgggaattcttagtcaagggtgctaaagtgcacacctcctgagcga 87
 Db 1453 AGATAGAAGATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1394
 QY 88 cctgtactggctatgttcgcgaagcctctactttgcagatcttagtaagcaagac 147
 Db 1393 RRR 1334
 QY 148 gcaacatctgttctgtgtaggaataacacctgaggtttgagttcttcttgcaggtt 207
 Db 1333 RRR 1274
 QY 208 gctttgagcagatgtatgaactaaagaaagccacgtccatggaacaaacctcaagagg 267
 Db 1273 RRR 1214
 QY 268 agcattcgaccagggattaggcaaacccgtgccactggagtcagagttgtgaggtgg 327
 Db 1213 RRR 1154
 QY 328 agggatgtctgttctcctgcggcagcggcggtgcccataatcattagagctctgaa 387
 Db 1153 RRR 1094
 QY 388 ctcatgtacaatgagtcacatcgtctacaacacatccagggtgaagatcgag 438
 Db 1093 RRR 1043

RESULT 7
 US-08-627-254C-26/c
 ; Sequence 26, Application US/08627254C
 ; Patent No. 5859229
 ; GENERAL INFORMATION:
 ; APPLICANT: Kniss, Douglas A.
 ; TITLE OF INVENTION: Eicosanoid Formation
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Calfee, Halter & Griwold LLP
 ; STREET: 800 Superior Avenue
 ; CITY: Cleveland
 ; STATE: Ohio
 ; COUNTRY: USA
 ; ZIP: 44114
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

128	cagatctagtaagaagagcgacaatactgttatgtgtaggaataaactctgttaggtt	187
Qy		
85	CATTTCGTAAAGAAAGAACAAAACATATTTCTTTAGAGGACAAATACTGGAGATT	944
Db		
188	tgatg	192
Qy		
945	TGATG	949
Db		



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 15:24:26 ; Search time 2934.5 Seconds
(without alignments)
1735.730 Million cell updates/sec

Title: US-09-236-995D-3

Perfect score: 474

Sequence: 1 aacaagatgctattatggca.....gtttccatcacaaaggtag 474

Scoring table:

IDENTITY_NUC

Gap 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	338.2	71.4	607	11	BF255013 HVSMF000
2	312	65.8	536	10	BE419432 WWS012.A1
3	299	63.1	661	10	AV834168 AV834168
4	269.2	56.8	369	10	BE420229 WWS04.C11
5	189	39.9	502	11	BG042229 su93c12.y
6	186.2	39.3	419	11	BE941860 EST421439
7	161.4	34.1	661	10	AV833893 AV833893
8	155.8	32.9	442	11	D42790 D42790 Rice
9	152.2	32.1	716	11	BI217470 602933675
10	150.6	31.8	587	11	BE911127 601662539
11	150.6	31.8	725	11	BI079902 602876096
12	142.6	30.1	571	10	BE304535 601105502

13	142.6	30.1	579	10	AA397988
14	142.6	30.1	669	10	AA401836
15	142.6	30.1	916	11	BI093436
16	142.6	30.1	952	11	BF793705
17	142.6	30.0	678	11	BF101770
18	141	29.7	709	10	BE382739
19	140.8	29.7	928	11	BG675913
20	136.8	28.9	487	10	AW144795
21	134.6	28.4	563	10	AA263755
22	134.6	28.4	728	10	AI455188
23	133.6	28.2	861	11	BG280821
24	130.6	27.6	975	11	BF026745
25	130	27.4	688	10	BE729084
26	127.6	26.9	560	11	BF042005
27	127.4	26.9	990	11	BG031594
28	125.4	26.5	425	11	BF769270
29	124.6	26.3	763	11	BG393399
30	124.6	26.3	911	11	BF793250
31	123.8	26.1	632	11	BG429402
32	123	25.9	308	10	BE123437
33	123	25.9	309	10	BE122975
34	122.6	25.9	771	11	BG751755
35	122.6	25.8	634	11	BI196670
36	122.2	25.7	499	10	BE488132
37	121.6	25.5	603	10	AI417842
38	121	25.4	617	10	AA212857
39	120.4	24.9	897	11	BE884815
40	117.8	24.7	742	11	BG303304
41	117	24.7	742	11	BG303304
42	116.2	24.5	465	10	AA608364
43	115.6	24.4	445	10	AA270953
44	115.2	24.3	593	11	W56030
45	114.8	24.2	438	10	BE334294

ALIGNMENTS

RESULT 1

BF255013

LOCUS 607 bp mRNA

DEFINITION HVSMF0005L18f Hordeum vulgare seedling root EST library HVCDNA0007 (etiolated and unstressed) Hordeum vulgare cdna clone

ACCESSION HVSMF0005L18f, mRNA sequence.

VERSION BF255013

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

REFERENCE 1 (bases 1 to 607)

AUTHORS Wing, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo, T., Sasaki, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and Wood, T.

TITLE Development of a genetically and physically anchored EST resource for barley genomics

JOURNAL Unpublished (2000)

COMMENT On Nov 16, 2000 this sequence version replaced gi:11184130.

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: AATTAACTCTACTAAAGG

High quality sequence stop: 551.

Location/Qualifiers

i..607

/organism="Hordeum vulgare"

TITLE
JOURNAL
COMMENT

Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2001)
Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kzsato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato.K., Saisbo.D., Takeda.K., Shini.T. and Kohara.Y. Direct
submission;
database: http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES
source

Location/Qualifiers
1. .661
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="rbags5k04"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare shoots germination"
/tissue_type="shoots"
/dev_stage="germination"
132 g 168 t
BASE COUNT 180 a 181 c 132 g 168 t
ORIGIN

Query Match 63.1%; Score 299; DB 10; Length 661;
Best Local Similarity 85.0%; Pred. No. 1.1e-79;
Matches 346; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 67 agaattgacacctctgagcactgttta-ctggctatatatttcggcaaaagccctactt 125
Db 661 AGGATACGACCTCTGTAGGCAACCCGTCACCGGCTATATGTTGGCAAAAGCCCTACTT 602
QY 126 tgcagatctagtaagcaagcgccacaataactgttatgtgataggaataatcctgtagg 185
Db 601 TGCAGATTTAGTAGCAAGAGTCACAGTATTGTTATGTGATAGAAAAATCCGACTGG 542
QY 186 ttgatgtctcttcttgaggtgcttcttaggagacatgtatgaactaaagaagccacgtc 245
Db 541 CTTGATGCTTCTTCTGAGGTGCTCTAGGAGACATTCATGAACCTGAAAAAAGCAAGCC 482
QY 246 catggacaaacctccaagaggaagcattcgaccagggattaggcaaaacccgtgcaact 305
Db 481 AATGGCAAAACCTCCAAAGGGGAAGCATTCGACCAAGGGCTTAGGCAAAACTGTGCCACT 422
QY 306 ggaatcagagtttggaagtgaggatgatgctgtagttccctgcgcaagccggtgcc 365
Db 421 AGAGTCGGACTTTGTTAAATGGAGGATGATGTCGTCGCTTGTGGCAAGCCAGTGC 362
QY 366 atcatcaattaggagctctgaactcatgtacaaatgagtaacatcgctacacacatccca 425
Db 361 AGCATCTATCAGGGCATCTGAGCTTCTGTACAAACGAGTATATAGTTACAAACACAGCTCA 302
QY 426 ggtgaagatcagttctgctgaggtgcttccatcaccaagagt 472
Db 301 GGTGAAGATGAGTCTCTTGTGTAAGTCAAAATTCCTGTCACAGCGTT 255

RESULT 4
BE420229 369 bp mRNA EST 24-JUL-2000
LOCUS
DEFINITION WMS04.C11R000101 ITEC WMS Wheat Scutellum Library Triticum aestivum
cDNA clone WMS04.C11, mRNA sequence.
ACCESSION BE420229
VERSION BE420229.1 GI:9418075
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticaceae; Triticum.
REFERENCE 1 (bases 1 to 369)
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier

S., Dubcovsky,J., Feullet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorells,M., Warburton,M. and Wenzel,G.
International Triticaceae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticaceae
Unpublished (2000)
Contact: Schuch W
Zeneca Wheat Improvement Centre, Norwich Research Park
Colney Lane, Norwich NR4 7UH UNITED KINGDOM
Tel: 44 1603 250 2600
Fax: 44 1603 250 699
Email: wolfgang.schuch@aguk.zeneca.com
International Triticaceae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.

FEATURES
source

Location/Qualifiers
1. .369
/organism="Triticum aestivum"
/cultivar="Novosibirskaya 67"
/db_xref="taxon:4565"
/clone="WMS04.C11"
/clone_lib="ITEC WMS Wheat Scutellum Library"
/tissue_type="scutellum callus"
/note="M13 Reverse sequencing primer used for 5' end of
clone."
78 c 101 g 87 t
BASE COUNT 103 a 78 c 101 g 87 t
ORIGIN

Query Match 56.8%; Score 269.2; DB 10; Length 369;
Best Local Similarity 85.2%; Pred. No. 9e-71;
Matches 312; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 64 ctaagaattgacacctctgagcactgttactggctatatgttcggcaaaagccctctac 123
Db 5 CGAGGATAGCACCTCCTGAGGCTCTCTGACAGGCTATATGTTGGCAAAAGCCCTCTAC 64
QY 124 ttgcagatctagtaagcaagcgccacaataactgttatgtgataggaataatcctgta 183
Db 65 TTTGAGATTTAGTAGCAAGAGTGCACAGTATTGTTATGTGGATAGAAAAATCCGACT 124
QY 184 ggttgatgctcttcttgaggttctttaggagacatgtatgaactaaagaagccagc 243
Db 125 GCCTTGATGCTTCTTCTGAGTTCCTCTAGGAGACATGCATGAACCTGAAAAAGCCACG 184
QY 244 tccatggacaaacctccaagaggggaagcattcgaccagggattaggcaaaacccgtgcca 303
Db 185 CCAATGGCAAAACCTCCAAAGAGGAAGCATTCGACCAAGGGCTTAGGAAAAACTGTGCCG 244
QY 304 ctggagtcagagtttggaagtgaggatgatgctgtagttccctgcgcaagccggtg 363
Db 245 CTAGAGTCGGAGTTTGTGAATGGAGGATGATATCGTCGCTTGTGGCAAGCCAGTG 304
QY 364 ccacatcaattaggagctctgaactcatgtacaaatgagtaacatcgctacacacatcc 423
Db 305 CCAGCATCTATCAGGGCATCTGAGCTTCTGT-CAACGAGTATATAGTCTACACACAGCC 363
QY 424 caggtg 429
Db 364 CAGGTG 369

RESULT 5
BG042229 502 bp mRNA EST 31-JUL-2001
LOCUS
DEFINITION su93c12.y1 Gm-c1055 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1055-2064 5', similar to TR:092P54 Q9ZP54 POLY(ADP-RIBOSE)
POLYMERASE ;, mRNA sequence.
ACCESSION BG042229
VERSION BG042229.1 GI:12488706
KEYWORDS EST.
SOURCE soybean.

ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 502)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 429 std Error: 0.00
High quality sequence stop: 419.

FEATURES

source
1..502
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl055-2064"
/clone_lib="Gm-cl055"
/tissue_type="Mature seed pods, greenhouse grown"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature seed pods of greenhouse grown plants prior to
senescence for the cultivar KPI. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments
were transformed into DH10B host cells (GibcoBRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker."

BASE COUNT 150 a 80 c 119 g 153 t
ORIGIN

Query Match 39.9%; Score 189; DB 11; Length 502;
Best Local Similarity 73.0%; Pred. No. 1.8e-46;
Matches 243; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 140 gcaagcgcaactactgtatgtggataggataatccctgtaggttgatgctcttt 199
DB 2 GCAAAAGTCTCAGTATTGCTTCACTGATAGAAAAATCCTGTGGCTTAATGCTTTGA 61
QY 200 ctgaggttgccttaggagcatgtatgaactaaagaaagccacgtccatggacaaacctc 259
DB 62 GTGAAGTTCGCTGCGTGGAAATGCTATGAGCTCAAGAAAGCTTAAGTATATATGATAAACCTC 121
QY 260 caagaggaagcattcgaccaggattagggcaaaacccgtgccactggagtcagatttg 319
DB 122 CGAAGAGGACACCTCTCTAAAGGACTGGGCAAGAAATGCCACAGGAATCGGAATATG 181
QY 320 tgaagtggagggatgatgtcgtagttccctgcggcaagccggtgcccattcaataggaa 379
DB 182 TAAAGTGGAGGGCAATGTCACCTGTTCTCTTGTGGCAACACAGTGCCATCAATGTCAGA 241

QY 380 gctctgaactcatgtacaatgagtagcatcgctctacacacatccaggtagaagtcagt 439
DB 242 GTTCTGAGCTCATGTCACATAGTATATTTTATATACTGCTCAAGTTAAGATGCAAT 301
QY 440 tcttgctgaaggtgcgtttccatcaccaagaggt 472
DB 302 TCTTATTGAAGTGAGGTGAGGTTTCATCACAAGAGAT 334
RESULT 6
BE941860
LOCUS BE941860
DEFINITION EST421439 MGHG Medicago truncatula cDNA clone PMGHG-6B18, mRNA
sequence.
ACCESSION BE941860
VERSION BE941860.1 GI:10519619
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 419)
AUTHORS Cote,F., Ojanen-Reuths,T., Hahn,M.G., VandenBosch,K., Hur,J.,
Beremand,P., Endre,G., Town,C.D., Bowman,C.L., Craven,M.B. and Cho
J., Fraser,C.M.
ESTs from seedling roots of Medicago truncatula after treatment
with beta glucan elicitor preparation from Phytophthora sojae
unpublished (2000)
CONTACT: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahncrcr.uga.edu
University of Georgia name: G269712e TIGR sequence name: MTAP09TK
More information is available at: http://chrysie.tamu.edu/medicago
Seq primer: SKmod (CTA gaa CTA gtg gat CC).

FEATURES

source
1..419
Location/Qualifiers
/organism="Medicago truncatula"
/cultivar="Al7"
/db_xref="taxon:3880"
/clone="PMGHG-6B18"
/clone_lib="MGHC"
/tissue_type="Roots from four day old seedlings"
/dev_stage="2 days after treatment with beta glucan
elicitor preparation from Phytophthora sojae"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-zap phage using Ex-assist
helper phage and propagated in XL0LR cells."

BASE COUNT 131 a 87 c 101 g 100 t
ORIGIN

Query Match 39.3%; Score 186.2; DB 11; Length 419;
Best Local Similarity 74.9%; Pred. No. 1.2e-45;
Matches 233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1 acaagatgctattatggcagcgttgcaagttgacaaatttctggaattcttagtcaa 60
DB 109 AACAGAATGCTCCTATGGCATGGTTCCTAGGCTGACCAACTTTGTGGCATCTTACCRA 168
QY 61 gggctaagaattgcacctctctgagcaccctgttactggctatatgttcggcaagcctc 120
DB 169 GGACTGAGAAATGCACCTCCCGAAGCCCTGCAACTGGTTATATGTTGGCCCAAGGATG 228

Db	132	TACTTTCCGCNACATGGTGTCCAAAAGTGCAAACTACTGCCACACATCTCAGGGAGACCCG	191
Qy	181	gtaggtttgatgcttcttcttaggttgcttttagggagacatgatgaactaagaagcc	240
Db	192	ATTGGCTTAATACTGCTGGGAGAGTGGCCCTTGGAAACATGTATGAATCTCAAGCATGCT	251
Qy	241	acgtccatggacaaacctcaagagaggaagcattcgaccaagggattaggacaaaccgctg	300
Db	252	TCACATATCAGCAAGTTACCCAGGGCCACACACAGTGTCAAAGGTTTGGAAAAACC---	308
Qy	301	ccactggagtcadaggtttgtgaagtggagggtatgtctgtagttcccttcgaggcaagccg	360
Db	309	ACCCTTGACCCTTCGGCCAGCATCACCTCGAGGGGTGTAGAGGTTCACCTGGGAACAGGG	368
Qy	361	gtgcacatcatcaattaggagctctgnaactcatgtacaatgagtcacatcgctctacaacaca	420
Db	369	ATCCCATCTGGTGTCAACGACACCTGCCTGCTGTATAAGTAGTACATTTGCTCTACGACATT	428
Qy	421	tcccaggtgaagatgcatgtcttctgtgaaggtgcgtttccatcacaca	469
Db	429	GCTCAGGTGAATCTCAATACCTGCTGAAACTCAAGTTCAATTTTAAAGA	477

[illegible]

CDNA Library Preparation: cDNA Libraries were prepared using the
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9129 row: n column: 09
High quality sequence stop: 559.
FEATURES
source
1. .587
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"

```

/ab_stages= 3 months, virgin
/ab_host="PH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: S
Site_2: NotI; Cloned unidirectionally. Primer: Oligo
Library construty by Life Technologies. Investigato
providing samples: Gilbert Smith, NIH"
BASE COUNT      149 a      163 c      150 g      125 t
ORIGIN
Query Match      31.8%; Score 150.6; DB 11; Length 587;
Best Local Similarity 59.1%; Pred. No. 8.1e-35;
Matches 277; Conservative 0; Mismatches 189; Indels 3; Gaps
w 1 aacaaatgctattatggacacgattcaagattacaaattttatgacaaattcttactcaa 60

```

```

Db 25 AACGGAGGCTGTGGCAGCGCTCAGGACCAACCACTTTGCTGGCATCTCTGTGCGAG 84
Qy 61 gggctgaagattgcacctctcctgagccacctgttactggtctatattgttgcgaagcctc 120
Db 85 GGTCTGGGATACCCCACTGAGCGCTGTGACAGGCTACATGTTTGGAAAGGGATC 144
Qy 121 taatttcagatctagtaagcaagagcgacataactgttattgtgdataggaataatcct 180
Db 145 TACTTTGCCACATGTGTCCAAAGTGAACACTACTGCCACACATCTCAGGAGACCGC 204
Qy 181 gtaggttgatgtttcttctgaagttgttttagggagacatgtatgaactaaagaagcc 240
Db 205 ATTGGCTTAATACTGCTGGGAGAGGTTGCCCTTGGAAACATGTATGAACCAAGCATGCT 264
Qy 241 acgtccatgcacaaacctccagaggaagcattcgaccaaaggtattagcgaagccgtg 300
Db 265 TCACATATCAGCAAGTTACCAAGGCGCAAGCAGTGTCAAGGTTTGGAAACACCC--- 321
Qy 301 ccactggatcgagttgtgaagtgaggatgtctgttagtccctgcggcaagccg 360
Db 322 ACCCTTGACCTTCGCCAGCATCACCTGGAGGGTGTAGAGGTTCCACTGGGAACAGGG 381
Qy 361 gtccatcatcaattaggagctctgaactcatgtacatagtaacatcgtctcacacaca 420
Db 382 ATCCCATCTGGGTCAACACACCTGCCTGCTGTATCATATGATGATACATTTCTACGACATT 441
Qy 421 tccaggtgaagatgcagttctgtctgtaagtgaggtgctgttagtccctgcggcaagca 469
Db 442 GCTCAGGTGAATCTCAATACCTGCTGAAACTCAAGTTCAATTTTAAAGA 490

RESULT 11
BI079902
LOCUS 602876096F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5007983 5',
DEFINITION BI079902 725 bp mRNA EST 20-JUN-2001
ACCESSION BI079902
VERSION BI079902.1 GI:14498232
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11051 row: o column: 24
High quality sequence stop: 664.
Location/Qualifiers
1. 725
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5007983"
/clone.lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

FEATURES
source
1. 725
/organism="Mus sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2988216"
/clone.lib="NIH_MGC_15"

```

```

BASE COUNT 177 a 204 c 188 g 156 t
ORIGIN
Query Match 31.8%; Score 150.6; DB 11; Length 725;
Best Local Similarity 59.4%; Pred. No. 8.7e-35;
Matches 274; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
Qy 9 gctattatggcagcggttcaaggttgacgaattttgtgggaatttcttagtcaagggctaa 68
Db 2 GCTGCTGTGGCAGCGCTCCAGGACCAACTTTGCTGGCATCTCTGTGCGAGGCTCTCCG 61
Qy 69 aattgcacctctcctgagccacctgttactggtctatattgttgcgaagccctctacttgc 128
Db 62 GATAGCCCCACCTGAAGCGCCTGTGCAGAGGCTACATGTTTGGGAAGGGATCTACTTGC 121
Qy 129 agatctagtaagcagcgcacatactgttattgttgataggaataatcctcttagttt 188
Db 122 CGACATGTTGTCCTCAAAAGTGAACACTACTGCCACACATCTCAGGAGACCCGATTGCTT 181
Qy 189 gatcttcttctcctgaggttctttaggagacatgtatgaactaaagaagccagctccat 248
Db 182 AATACGTGCTGGAGAGGTTGCCCTTGGAAACATGTATGAACCTCAAGCATGCTTCACATAT 241
Qy 249 ggacaaacctccagaggggaagcattcgaccaaaggtattagcgaagccgtgccaatgga 308
Db 242 CAGCAAGTTACCAAGGCGCAAGCAGTGTCAAGGTTTGGGAACACCC---ACCCCTGA 298
Qy 309 gtccaggtttgtgaagtgaggatgtatgtctgttagtccctgcggcaagccgtgccaatc 368
Db 299 CCCTTCGGGCCAGCATCACCTTCGAGGGTGTAGAGGTTCCACTGGGAACAGGATCCCATC 358
Qy 369 atcaattaggagctctgaaactcatgtacatagtcacatcgtctcacacacatcccaagt 428
Db 359 TGGTGTCAACGACACCTGCCCTGCTGTATATGAGTACATGTTCTTACGACATGCTCAGGT 418
Qy 429 gaagatcgagttcttctggaaggtgcgttttccatcacacaaga 469
Db 419 GAATCTCAAAATACCTGCTGAAACTCAAGTTCAATCTTTAAGA 459

RESULT 12
BI0304535
LOCUS 601105502F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2988216 5',
DEFINITION BI0304535 571 bp mRNA EST 13-JUL-2000
ACCESSION BI0304535
VERSION BI0304535.1 GI:9175756
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 571)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM77 row: c column: 01
High quality sequence stop: 567.
Location/Qualifiers
1. 571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2988216"
/clone.lib="NIH_MGC_15"

FEATURES
source
1. 571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2988216"
/clone.lib="NIH_MGC_15"

```

```

/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

```

```

BASE COUNT      147 a      140 c      144 g      144 t
ORIGIN

```

```

Query Match      30.1%; Score 142.6; DB 10; Length 571;
Best Local Similarity 58.0%; Pred. No. 2.1e-32;
Matches 272; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 1 acaagatgctattatggcagcgttcaaggttgacgaatttgggaattcttagtcaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40 AACCGAAGATTGCTGTGGCAGGGTCCAGGACCACCACTTTGCTGGGATCCTGCCAG 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 gggctaagaattgacacctctgagcagcctgttactggtctatatgttcggcaaggcctc 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100 GCTCTTCGGATAGCCCGCTCGAAGCGCCGCTGACAGGCTACATGTTTGGTAAAGGGATC 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 tactttgagatctagtgaagcaagagcgcacaaatactttatgttgataggataatcct 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160 TATTTTCGTGACATGGTCTCCAAAGAGTGCCAACTACTGCCATAGCTCTCAGGGAGACCCA 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 gtaggttgatgctctcttctgaggttcttttaggagacatgtatgaactaaagaagcc 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 ATAGGCTTAATCCTGTGGGAGAGTGGCCCTTGGAAACATGTATGAACACACGCT 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 acgtccatggacaaacctccaagaggaagcattcgacaaaggattaggcaaaaccgct 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 TCACATATCAGCAAGTTACCCAAAGGCAAGCAGACAGTGTCAAAGGTTTGGGCAAAATACC 339
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 ccactggagtcagagttgtgaagtggagggatgatgcgttagttccctcgcaagccg 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 CCGTATCCCTTCAG---CTAACATTAGTCTGGATGTTGAGACGTTCTCTTGGGACCGGG 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 gtgcatcatcaattaggagctctgaactcatgtacaatgagatcatcgtctacaacaca 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 ATTTCATCTGGTGTGATGACACCTCTCTACTATATACAGGATACATGCTATGATATT 456
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 tccaggtgaaagtcagttcttctgtgaggtgcgctttccatcacaaga 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 GCTCAGGTAATCTGAAGTATCTGCTGAAACTGAAATCAATTTTAAAGA 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
AA397988      579 bp      mRNA      EST      12-AUG-1997
LOCUS      zt86c12.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:729238
DEFINITION      5' similar to gb:J03473 NAD(+) ADP-RIBOSYLTRANSFERASE (HUMAN);
                mRNA sequence.
ACCESSION      AA397988
VERSION      AA397988.1 GI:2050655
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 579)
AUTHORS      Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
                Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B.,
                Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
                ,T., Waterston,R. and Wilson,R.
                WashU-Merck EST Project 1997
                Unpublished (1997)
                Contact: Wilson RK
                Washington University School of Medicine

```

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estwatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1320 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 427.

```

```

FEATURES
    source
        1..579
            /organism="Homo sapiens"
            /db_xref="GDB:5925767"
            /db_xref="taxon:9606"
            /clone="IMAGE:729238"
            /clone_lib="Soares_testis_NHT"
            /sex="male"
            /lab_host="DH10B"
            /note="Vector: pT73D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech Laboratories
            , Inc., and primed with a Not I - oligo(dT) primer [5',
            TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT73 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaudo. "

```

```

BASE COUNT      148 a      138 c      144 g      149 t
ORIGIN

```

```

Query Match      30.1%; Score 142.6; DB 10; Length 579;
Best Local Similarity 58.0%; Pred. No. 2.1e-32;
Matches 272; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 1 acaagatgctattatggcagcgttcaaggttgacgaatttgggaattcttagtcaa 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43 AACCGAAGATTGCTGTGGCAGGGTCCAGGACCACCACTTTGCTGGGATCCTGCCAG 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 gggctaagaattgacacctctgagcagcctgttactggtctatatgttcggcaaggcctc 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 GGTCTTCGGATAGCCCGCTCGAAGCGCCGCTGACAGGCTACATGTTTGGTAAAGGGATC 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 tactttgagatctagtgaagcagcgcacaaatactgttattgtaggaataatcct 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 TATTTTCGTGACATGGTCTCCAAAGAGTGCCAACTACTGCCATAGCTCTCAGGGAGACCCA 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 gtaggttgatgctcttctgtgaggtgttccttagggagacatgtatgaactaaagaagcc 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 ATAGGCTTAATCTGTTGGGAGAGTTCCTCTGGAACATGATGAACCTGAAGCAGCT 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 acgtccatggacaaacctccaagaggaagcattcgacaaaggattaggcaaaaccgct 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 TCACATATCAGCAAGTTACCCAAAGGCAAGCAGACAGTGTCAAAGGTTTGGGCAAAATACC 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 ccactggagtcagagttgtgaagtggagggatgatgcgttagttccctcgcaagccg 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 CCGTATCCCTTCAG---CTAACATTAGTCTGGATGTTGAGACGTTCTCTTGGGACCGGG 399
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 gtgcatcatcaattaggagctctgaactcatgtacaatgagatcatcgtctacaacaca 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 ATTTCATCTGGTGTGAAAGACACCTCTCTACTATATTAACGAGTACATGCTATGATATT 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 tccaggtgaaagtcagttcttctgtgaggtgcgctttccatcacaaga 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 GCTCAGGTAATCTGAAGTATCTGCTGAAACTGAAATCAATTTTAAAGA 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
AA401836      669 bp      mRNA      EST      16-MAY-1997
LOCUS      zv66b12.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION

```


Db 365 ATAGGCTTAATCCTGTTGGGAGAAAGTTGCCCTTGGAAACATGTATGAACCTGAAGCAGCT 424
 Qy 241 acgtccatggacaaacctccaagagggaagcattcgaccacaaaggattaggcaaaacccgtg 300
 Db 425 TCACATATCAGCAAGTTACCCCAAGGGCAAGCACAGTGTCAAAGGTTTGGGCAAAACTACC 484
 Qy 301 ccaactggagtcagagattttgtgaagtggaggagatgatgtcgtagttccctgcgcaagccg 360
 Db 485 CCTGATCCTTCAG---CTAACATTAGTCTGGATGGTGTAGACGTTCCTCTTGGGACCCGGG 541
 Qy 361 gtgcacatcatcaataggagctctgaactcatgtacaaatgagtagacatcgtctacaacaca 420
 Db 542 ATTTTCATCTGGTGTGAATGACACACCTCTCTACTATATAACGAGTACATTGTCTATGATATT 601
 Qy 421 tcccagggtgaagatgcagttcttgcgaaggtgcggtttccatcaacaaga 469
 Db 602 GCTCAGGTAATCTGAAGTATCTGCTGAACCTGAAATTCAAATTTAAGA 650

Search completed: March 7, 2002, 15:24:30
 Job time: 3340 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 14:32:50 ; Search time 4873.04 Seconds
(without alignments)
9983.538 Million cell updates/sec

Title: US-09-236-995D-1
Perfect score: 2949
Sequence: 1 atggcgccgcgcaaaaggc.....gtttccatcacagaaggttag 2949

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.ov.*
- 5: gb.pat.*
- 6: gb.ph.*
- 7: gb.pl.*
- 8: gb.pr.*
- 9: gb.ro.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.ba.*
- 15: em.fun.*
- 16: em.in.*
- 17: em.in.*
- 18: em.in.*
- 19: em.in.*
- 20: em.ov.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo_hum.*
- 31: em.htgo_inv.*
- 32: em.htgo_rod.*
- 33: em.htg_hum.*
- 34: em.htg_inv.*
- 35: em.htg_rod.*
- 36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2797.6	94.9	3285	8	AF093627	AF093627 Zea mays
2	2720	92.2	3211	8	ZMPARP2	AT222589 Zea mays
3	1060.2	36.0	3187	8	ATH131705	AT131705 Arabidops
4	249.4	8.5	3036	5	GGPADPRP	X52690 Chicken mRN
5	245.8	8.3	3589	4	BOVPAS	D90073 Bovine poly
6	239.2	8.1	3718	5	XELPADPRP	D14667 Frog mRNA f
7	238.2	8.1	3617	5	XLPARPG	Z12139 X.laavis PA
8	224.6	7.6	3640	9	HUMPOLP	M18112 Human poly(
9	224.6	7.6	3792	6	A52134	A52134 Sequence 1
10	224.6	7.6	3795	9	HUMRISAD	J03473 Human poly(
11	223	7.6	1771	9	HUMADPPO	M17081 Human place
12	222.2	7.5	3047	10	AF168781	AF168781 Crictetulu
13	219.8	7.5	1939	10	RNPARP2	X65497 R.norvegicu
14	219.8	7.5	3045	6	AX058340	AX058340 Sequence
15	219.8	7.5	3045	6	AX062277	AX062277 Sequence
16	219.8	7.5	3660	9	HUMPOL	M32721 Human poly(
17	219.8	7.5	3747	6	I14359	I14359 Sequence 5
18	217.4	7.4	3128	10	RNU94340	U94340 Rattus norv
19	217.2	7.4	3845	10	BC012041	BC012041 Mus muscu
20	216	7.3	1732	10	AF126717	AF126717 Mus muscu
21	212.4	7.2	3172	10	MMADPRP	X14206 Mouse mRNA
22	205.2	7.0	3200	6	AX058362	AX058362 Sequence
23	204.4	6.9	2147	8	ATPARP	Z48243 A.thaliana
24	204.2	6.9	3463	3	SPEPADPRP	D16482 Sarcophaga
25	200.2	6.8	2295	8	ZMPARP1	AJ222588 Zea mays
26	191.8	6.5	1754	9	HS236912	AJ236912 Homo sapi
27	191.8	6.5	1814	6	AX058317	AX058317 Sequence
28	191.8	6.5	1843	6	AX008892	AX008892 Sequence
29	191.8	6.5	1874	6	AX058374	AX058374 Sequence
30	191.6	6.5	3080	3	DROADPRP	D13806 Fruit fly m
31	190.2	6.4	1797	9	HS236876	AJ236876 Homo sapi
32	190.2	6.4	1910	9	AF085734	AF085734 Homo sapi
33	186.2	6.3	121158	2	AP003889	AP003889 Oryza sat
34	185.4	6.3	1980	9	AK001980	AK001980 Homo sapi
35	183.2	6.2	1822	3	DMPARP5	AF051548 Drosophil
36	183.2	6.2	10988	2	AC014613	AC014613 Drosophil
37	183.2	6.2	16864	3	AF002892	AF002892 Drosophil
38	179.8	6.1	1707	6	AX058338	AX058338 Sequence
39	179.8	6.1	1707	10	MMU7780	AJ007780 Mus muscu
40	179.8	6.1	1766	10	AF072521	AF072521 Mus muscu
41	175.8	6.0	79663	8	AC006593	AC006593 Arabidops
42	171	5.8	864	6	AX058336	AX058336 Sequence
43	163.8	5.6	687	6	AX058382	AX058382 Sequence
44	163.8	5.6	738	6	AX058380	AX058380 Sequence
45	163.2	5.5	5787	9	GORADPRB	L24094 Gorilla gor

ALIGNMENTS

RESULT 1

AF093627

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AF093627

Zea mays

poly(ADP)-ribose polymerase (PARP1) mRNA, complete cds.

AF093627

AF093627.1

GI:3928870

Zea mays.

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 3285)

Mahajan,P.B. and Zuo,Z.

Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase

Plant Physiol. 118 (3), 895-905 (1998)

99026291

2 (bases 1 to 3285)

Mahajan,P.B. and Zuo,Z.

Direct Submission

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

[illegible]

Qy	2638	gtgataggaaataacctgttaggtttgatgcttctttcttgaggttgcttttagggagacaatg	2697
Db	2731	GTGATAGGAATAACTCTGTAGTTAGTTGATGCTCTTTCTGAGGTGCTTTTAGGAGACATG	2790
Qy	2698	tatgaactaaagaagaacacgtccatggacaaaacctccaagaggaagcattcgaccaag	2757
Db	2791	TATGAACATAAGAAGACCGTCGATGACAAACCTCCAAAGAGGAGCAATTCGACCAAG	2850
Qy	2758	ggattaggcaaaaaccgtgccactgagtcagagtttgtgaagtggaggatgatgtcgta	2817
Db	2851	GGATTAGCAAAACCGTCCCACTGGAGTCAGAGTTTGTGAAGTGGAGGATGATGTCGTA	2910
Qy	2818	gttccttcggccaagccggtgccatcatcaattaggagactctgaactcatatatacaatgag	2877
Db	2911	GTTCCTCGCGCAACCGGTGCCATCATCAATTAGGAGCTCTGAACCTCATGTACAAATGAG	2970
Qy	2878	tacatcgctctacaacacatccacagtggaagatgcagttcttctgtaaggtgcgtttccat	2937
Db	2971	TACATCGTCTACAACACATCCACAGTGAAGATGCAGTTCTTGCTGAAGGTGCGTTTCCAT	3030
Qy	2938	cacaagagtgtag	2949
Db	3031	CACAAGAGGTAG	3042
RESULT 2			
ZMPARP2			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
1 (bases 1 to 3211)			
BABYCHUK E., Cottrill, P., Storozhenko, S., Fuangthong, M., O'Farrell, M., Van Montagu, M., Inze, D. and Kushnir, S.			
Higher plants possess two poly(ADP-ribose) polymerases			
Unpublished			
2 (bases 1 to 3211)			
Kushnir, S.			
Direct Submission			
Submitted (14-NOV-1997) Kushnir S., VIB, Dep.Genetics, Ledeganckstraat 35, Gent, B9000, Belgium			
Location/Qualifiers			
1. .3211			
/organism="Zea mays"			
/db_xref="taxon:4577"			
113. .3022			
/gene="PARP"			
113. .3022			
/gene="PARP"			
/function="secondary protein modification"			
/codon_start=1			
/product="poly(ADP-ribose) polymerase"			
/protein_id="CAAL0889.1"			
/db_xref="GI:2632129"			
/db_xref="SPTREMBL:O24570"			
/translation="MAAPPKAAEYAKSGRACKSCRSPIADQLRLGKMYQASOF GMPMNHNSVDVEGIDALRWDDQEKIRNYGSSAGSSPTAAPEPKETIELAPSA TSCRCRSKITGTGSRVRLSAKLSGEPKGIPIWHANCFFEVSPATVEKESGMDTSLSD DKRTMLDLVKDGVNNEKNGSKRKSENDIDSYSARLDETSCTGVNRKNGOLVDR GNSTSADLIQLKLEQSDTLWKLDGLKTHVAALRDLMLANGQDTSGPERHLLDR ADGMIFGALGPCVPCANGWYYNGPOCSGNSWSEKCYTSATPEVVRKKWQIPQIH KNDYLNKWFKSQKPEVLPDPSKTSKATORTLSLSKGLDKLRFVSVVGQSK ANNEWLEKLNKALNYARVKDIDLACGELDENAEVVKARLLKIPIVREGYIG CVYKNMLPFDLYKLNALNLESGSKTIVVKGRSAVHESSGLQDTHAHLEDKSYI ATLNMSDIALGNVSYVLOIIEQDDGSECVYFRWGRVSGEIGGQKLENSKTEAL EPRKLFETKNSWEAECKTNRKQPGFYPLDVGYVKAPKAPRDLSEMKSSLAJ LAFIWKMIFNVTYIRAAWPEFLNMSEMPLGKLSKENIEKGFALTPEIQNLKDTAT			
FEATURES			
source			
gene			
CDS			

ALAVRESLIVASRNFFLLIPSIHPIIIRDEDDLMIRAKMLLEALQDIEIASKIVGFDS
 DSESLDDKYNKLDIIPPLAHDIEDYKLIHOYLLNTHAPTHKDWSELEVRVSLDRD
 GBLNKYSRVKNLHNMKMLLWHSRLTNEVGLISGLRLTAPPEAVPYSTMGEKGLIFAD
 LVSKSAOYCYVDRNNPVCMLLSLWALGDMYELKATYSMDKPPPKKHSTKGLKTVPL
 ESEFKWRRDDVVPCGKVPVPSIRSSLEMYNEYIYVNTSVQKMOFLKVRPHHR*

BASE COUNT 968 a 604 c 813 g 826 t
 ORIGIN

Query Match 92.2%; Score 2720; DB 8; Length 3211;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 2869; Conservative 22; Mismatches 16; Indels 45; Gaps 11;

QY 1 atggcgccgcccacaaagcgtggaagcgagtgatgccaagctggcgccctcgagc 60
 DB 113 ATGGCGGCGCCCAAAAGCGTGAAGCGGAGTATGCCAAGCTGCGGGCGCCCTCGGCG 172
 QY 61 aagtcacgcccgtccctatgcgccaagaccagctccctcttgccaagatggttcagcgc 120
 DB 173 AAGTCACGCGCGTCCCTATCGCCAAGACCAAGCTCCGCTTGGCAAGATGCTTCAGGCG 232
 QY 121 tcacagttcgcagcgttcctcatgcgagtggaaccatgcaggtgcacatcttcagcaagaag 180
 DB 233 TCACAGTTCGAGCGGCTTCATGCCGATGTGAACCATGCCAG----- 273
 QY 181 aaccagataaaatccgttgacgatgttgaaggatagatgcacttagatgggatgacaa 240
 DB 274 -----CGTTGACGATGTTGAAGGATAGATGACATAGATGGGATGATCA 319
 QY 241 gagaagatacagaactacgttggagtgctcagctggttacaagtctcagctgcctc 300
 DB 320 GAGAAGATACGAACACTACGTTGGAGTGCTCAGCTGGTACAAGTCTCAGAGCTGCTCT 379
 QY 301 cctgagaatgtacaatgagatgctccatctccgctaccttcacatagacgatgcagt 360
 DB 380 CCGTGAAGTGTACAAATGAGATGCTCCATCTGCCGCTACCTCATGTAGACGATGCAGT 439
 QY 361 gaaagattacaaagagtcggtccgtcttccagctaaagctgagagtgaggtcccaag 420
 DB 440 GAAAGATGTACAAAGGATCGTCCGCTTTTCAGCTAAGCTTGAGAGTGGAAGTCCCAAG 499
 QY 421 ggtatacagtgatcagcaactgttctttgaggtatcccgctgcaactgtgtgag 480
 DB 500 GGTATACCATGGTATCATGCAACTGTTCTTTGAGGTATCCCGCTGCAACTGTGTAG 559
 QY 481 aagttctcagcgtggatacttttccgatgaggaagaagaccatgctcagcttgtt 540
 DB 560 AAGTTCTCAGGCTGGGATCTTTGTCGATGAGGATGAAGAACCATGCTCGATCTGTGT 619
 QY 541 aaaaaagattgtggaacaatgaacaaataaagggttcccaagcgcaagaagaagtgaat 600
 DB 620 AAAAAAGATGTGGCAACAATGAACAAATAAAGGTTCCTCAAGCGCAAGAAAGTGAAT 679
 QY 601 gatattgatagctacaaatccgcaggttagatgaaggtacatctgaaggtacagtcga 660
 DB 680 GATATTGATAGCTACAAATCCGCGAGGTTAGATGAAGTACATCTGAAGGTACAGTSCGA 739
 QY 661 acaaaagggcaactttagaccacagctggttcccaatcactagttcagctgatctccaacta 720
 DB 740 AACAAAGGCAACTGTGATACCCACGCTGTTCCAAATAGTTCACGCTGATATCCCACTA 799
 QY 721 aagcttaagagagaagtgacacactttggaagttaaaagtgagacttaagactcatgta 780
 DB 800 AAGCTTAAGGAGCAAGTGTACACACTTTGGAAGTTAAAGGATGAGCTTAAGACTCATGTA 859
 QY 781 tcggctgctgaatgaaggatagcttgagcttaagcttaagcttaagcttaagcttaagct 840
 DB 860 TCGGCTGCTGAATTAAGGATATGCTTTGAGGCTTAATGGCGAGGATATACATCAGACAGAA 919
 QY 841 aggcacctattgatgcgtgtgcggatggaatgctatttggagcgtggtccttgccca 900
 DB 920 AGGCACCTATTGATGCTGTGCGGATGAATGATATTATGGAGCGCTGGGCTCCTTGCCA 979

QY 901 gtctgtgctaatggcagtcagtactattataatgggtcagtcaccaatgcagtgtaagtgtca 960
 DB 980 GTCTGTGCTAATGGCATGTACTATTATAATGGTCAAGTACCAATGCAGTGGTAATGTGTCA 1039
 QY 961 gagtgggtccaaagtgtacatactctgcacagaacctctccgcttaagaagaagtgagcaa 1020
 DB 1040 GAGTGGTCCAAAGTGTACATCTGCCCACAGAACCTCTCCGCGTTAAGAAGAAGTGGCAA 1099
 QY 1021 attccacatggacaagaagattacattatgaagtgggttcaaaatctcaaaagggttaag 1080
 DB 1100 ATTCCACATGCAACAAAGAATGATTACCTTATGAAGTGGTTCAAATCTCAAAAGGTTAAG 1159
 QY 1081 aaaccaagaggggttcttcacacaaatgcaccttgagaatctggaagtaagaacaactcag 1140
 DB 1160 AAACCAAGAGAGGGTCTTCCACCAATGTCACTTCCACCTGAGAAATCTGGAAGTAAAGCAACTCAG 1219
 QY 1141 agaacaatcattgctgtcttcttaaaaggttgataaataaaggtttctctgtgtgagacaa 1200
 DB 1220 AGAACATCATTTGCTGCTCTTAAGGGTGGATTAATTAAGTGGTTCCTGTTGTAGGACAA 1279
 QY 1201 tcaaaagaaagcagaatgagtgattgagaagctcaaaactgtggtgccaacttctat 1260
 DB 1280 TCAAAAGAAAGCAGCAATGATGATGAGAGCTCAAACTTCTGCTGCCAACTTCTAT 1339
 QY 1261 gccaggttctcaaaatattgatttttaattgaatgagtggtgagctgcacaaatgaaa 1320
 DB 1340 GCCAGGGTGTCAAAAGATATTGATTGTTAAATGCGATGTTGGTGCAGCTCGACAAATGAAAT 1399
 QY 1321 gctgaagtcaggaaagcaggagcgtgaagatacacaattgtaaggagaggttaccattgga 1380
 DB 1400 GCTGAAGTCAGGAACCAAGGAGGCTGAAGATACCAATTTGTAAGGGAGGGTTACATTGGA 1459
 QY 1381 gaatgtgtt-aaagaaacaaatgctgcaattgtattgtataaaact--ggaatgcctta 1437
 DB 1460 GAATGTGTTAAAGAAACAAATGCTGCCATTTGATTTCTATAAACTAGAGAATGCCTTA 1519
 QY 1438 gagtccctcaaaagcmgtactctcactgttaagtttaagggcggaagtgctgttcata 1497
 DB 1520 GAGTCTCTAAAGGCGAGTACTGCTACTGTTAAAGTTAAGGGCGCAAGTCTGCTTC-ATG 1577
 QY 1498 agtccctcyggttgcagaataactgctcactctcacttcttcttcttcttcttcttcttctt 1557
 DB 1578 AGTCTCTGTTTGAAGA-TACTGCTCACAAT-CTTGAAGATGGGAAAGCATATACAA 1635
 QY 1558 tgcacmcttaacacatgtctgacccctgncacnagggtgtgacaggtctactatgtactccag 1617
 DB 1636 TGCAACCTTTAAACATG-TCTGACCTGGCAGTACTAGGTGCTGAACA-GCTACTATGTACTCCAG 1693
 QY 1618 atcattgaacagagtgatgggtcgtgagtgctacgtatttctgaagtgagggaggttggg 1677
 DB 1694 ATCATTTGAACAGAGTATGGGTCTGAGTGTCTGAGTATTTTCGTTAAGTGGGAGGGTGGG 1753
 QY 1678 agtgagaaaaattggaggggcaaaactgagagagatgtcaaaactgagcgaactcaagaa 1737
 DB 1754 AGTGAGAAAAATTGGAGGGCAAAACTGGAGGAGATGTCAAAACTGAGGCAATCAAGAA 1813
 QY 1738 ttcaaaagattatttcttgagaagactggaactcagggagccttgggaagctgaaaccc 1797
 DB 1814 TTCAAAAGATTATTCTTTGAGAAGACTCGAAACTCATGGGAGCTTGGGAATGTAAACCC 1873
 QY 1798 aatttctggaagcagcctgggagattttccaccacttgatgttcttcttcttcttcttcttct 1857
 DB 1874 AATTTTCGGAAGCAGCTTGGGAGATTTTACCACCTTGTATGTTGATTTGTTGTTAAGAAA 1933
 QY 1858 gcacaaacacggaagatatcagtgaaatgaaaagtctcttctctctctctctctctctctct 1917
 DB 1934 GCACCAAAACGGAAGATATCAGTGAATGAAAGTCTCTTCTGCTCTCAATTCTCTAGAA 1993
 QY 1918 ctcatgaagatgcttttcaatgtgagacatatagacgtctatgatggaatttgaawt 1977
 DB 1994 CTCATGAAGATGCTTTCAATGTGGAGACATATAGAGCTGCTATGATGGAAATTTGAAAT 2053
 QY 1978 aatgtgcagaatgcctctcttcttcttcttcttcttcttcttcttcttcttcttcttct 2037

[illegible]

RESULT 3

ATH131705

LOCUS

DEFINITION

ACCESSION

ATH131705 3187 bp mRNA PLN 18-DEC-1998
Arabidopsis thaliana mRNA for poly(ADP-ribose) polymerase.

VERSION	AJ131705.1	GI:4038490
KEYWORDS	NAD(+) ADP-ribosyltransferase; parp-1 gene; poly(ADP-ribose) polymerase.	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi	
AUTHORS	1 (bases 1 to 3187)	
TITLE	Kazmaier,M.	
JOURNAL	Direct Submission	
FEATURES	Submitted (17-DEC-1998) Kazmaier M., Departement D'Eco	
AUTHORS	Vegetale Et De Microbiologie, Commissariat A L'energie Atomique, CEA Cadarache, Lab. De Radiobiol. Vegetale, Bat. 185, 13108 St.	
JOURNAL	Paul-les-Durance, FRANCE	
FEATURES	2 (bases 1 to 3187)	
AUTHORS	Doucet-chabeaud,G. and Kazmaier,M.	
JOURNAL	Unpublished	
FEATURES	Location/Qualifiers	
source	1..3187	
gene	/organism="Arabidopsis thaliana"	
CDS	/cultivar="landsberg erecta"	
	/db_xref="taxon:3702"	
	11..2963	
	/gene="parp-1"	
	11..2963	
	/gene="parp-1"	
	/standard_name="NAD(+) ADP-ribosyltransferase"	
	/EC_number="2.4.2.30"	
	/function="ADP-ribose polymer synthesis"	
	/note="mol. weight: 111 kDa"	
	/codon_start=1	
	/evidence=experimental	
	/product="poly(ADP-ribose) polymerase"	
	/protein_id="CA10482.1"	
	/db_xref="GI:4038491"	
	/translation="MASPKPWRAEYAKSSRSSCKTKSVINKENFRLKLVOSTHFD	
	GIMPNWASCIILKTKWQDKISVDVDEGIESLRWDEKIKHYVESGAGNSTSTGTGS	
	TSSANNAKLEYIEQYSQRACRCKSEKILKEVRIFSPKPGNGKGLMWHKACF	
	LEMSSTELKSSGWSIPDSQDEALLPLVKALPAKKTAEARQNTSRAGTKRKND	
	SVNDEKSLKASDFMSTGALQPCCKEMEQAOTKELMDKLKKYVTSAEELRM	
	EVNEQTRSEKLLDRKCDGMGFWPLKNCSPGSHLSPSGGLYRCHGYLSWSKSCSH	
	TLDPRIRIGKWKIPUETENQFLFKNKSKSVKPKRILRPVLSETSGSQSGSKDRAT	
	SSRSERLADKVSIAQNTQRPWKARIEAGAEFANVKKGTSLVVCGLTDIARE	
	MKRARKMVAIVREDFLVDCFKQKRLPDFKYIKEDTSLSLVTKVYKGRSAVHAESGL	
	QEHCHTLDGNSYNTILDSMDLSTGINSYIIQIIQEDGSDCYFRKWRGVNKKI	
	GNGNEMSKSNAPAEHFELKLEKTTGWESQKTFNQPCGKFLPLDIDYGVNKKI	
	AKKEPQTSNLAPELIMKMLFDVETYSAMWEFIINMSEMPGLKSLKHNKQKGE	
	ALTEIORLLTESDPQTMKSLVLDASNREFTMIPSIHPHIIDEDDFKSKVMLEAL	
	QDIEIASRVGDFDVSTEDLDDKKLKHCDISLPHDSEDYRLIEKLYLNTHTAPTTPE	
	WSLEIEFVALEREHGFYAPHREKLGKMLHGHGSRFTNFVGLINQGLQRLTAPTPE	
	ATGFMFGKGIYFADLVSKAQYCTCKNPKVLMLLSEVALGELHETLKAKYMDKRPQ	
	KGHTKGLGKGVQDSEFAKRWGDVTVPCGKPVSSKVASELMYNEYIVDTAQVVKLQ	
	FLKLKVRFKHKKR"	
mat_peptide	14..2959	
	/gene="parp-1"	
	/EC_number="2.4.2.30"	
	/product="poly(ADP-ribose) polymerase"	
BASE COUNT	1044 a	577 c 743 g 823 t
ORIGIN		

Db 11 ATGGCAAGCCACATATAGCGTGGAGGGCGAGTATGCAAGCTCTCAGGCTTTCATGT 70

Qy 61 aagtcatgccggtccctctatcgcaaggaccagctccgcttggcaagatggttcaggcg 120

Db 71 AAAATCTGCAAGTCCGCTCATTAACACAGGAGAACTTCGCTTGGAAAGCTGGTTCAATCT 130

|||||
Db 2212 CCTCATGACTTGGGATGAAGAACCCACCTCTCTCAGTAACTTGGAAATACATTCAGGCT 2271
Qy 2194 aagggcgaatgcttgaagctctgcaggatattgaaattgcttcaaa-----agatagtt 2247
Db 2272 AAGTGCAGATGTTGGACAACCTGCTCGATATGAGGTTGCTTACAGCTTCTCAGAGGT 2331
Qy 2248 ggtctcgatagcagcagtgatgaatctcttgatgataaataatgaacattcactgtgac 2307
Db 2332 GGAATGAGATGGAGACAAAGACCCCAATTTGATATCAATTATGAAAGCTCCGAACTGAT 2391
Qy 2308 atcacccgcgtgctccagatagtgaaattacaaagtttaattgagcagatctctcctcaac 2367
Db 2392 ATTAAGTGGTTGACAAAGATTTCAGAACGACCAAGATTATTAACAAATACGTGAAAAAT 2451
Qy 2368 acacatgctctactcaacaggactggtcgctggaactggaggaggttttttcaacttgat 2427
Db 2452 ACTCATGCTGCTACCTACCAATGCATATGACCTCAAGTTGGAATCTTCAGGATTCAA 2511
Qy 2428 cgagatgagaaacttaataagctcaagatataaaataatctgcatacaacagatgcta 2487
Db 2512 CTTGAAGGAGAGATCAGCTTACAGCCATTT---AAGCAGCTTCATAATCGCCAGCTG 2568
Qy 2488 ttatggcgcgttcaaggttgacgaattttgtgggaattttagtcaagggtcagaatt 2547
Db 2569 CTGTGCGACGGTTCCGCCACCAACCACTCGCTGTATCTCTCACAGGGTCTCCGGATA 2628
Qy 2548 gaactctgagcagcactgttactgcttatgttgcgaagcctcactcttgatgagat 2607
Db 2629 GCTCCCTGGAAGCTCTGTGACCGGCTACATGTTTGGGAAGGCACTCTACTTTCGAGAT 2688
Qy 2608 ctagtaagcaagcgcacaaactctgttatgtggataggaataatcctgtaggtttgatg 2667
Db 2689 ATGGTTTCCAAGAGTGCCAACTACTGTCTCACATCTCAAGCTGATCCAAATTTGGTTAATA 2748
Qy 2668 ctcttcttgaggttctttaggagacatgtatgaactaaagaagccacatccatcgagac 2727
Db 2749 CTACTGGGAGAAAGTTCCTCGGAAATATGTATGAGCTTAAAGATGCTTCTCACATAACA 2808
Qy 2728 aaactccaagagggaagcattcgaccagaggttaggcaaaaacggtgcccactgagatca 2787
Db 2809 AAATTTGCCCAAGGAAAACATAGTGTGAAGGCTTGGGCAAAACTG---CACCTGATCCC 2865
Qy 2788 gaggttggaagtggaggtgatgctgtagtccctcggaagcgggtccatcatca 2847
Db 2866 ACAGCCACTACAAACCTTGTAGTGTAGAAAGTTCCTTAGGAAATGGGATCTCAACAGGA 2925
Qy 2848 attagagctctgaactcatgtacaaatgagtagatcgtctacacacacatccaggtgaag 2907
Db 2926 ATTAATGATACTTGTCTGCTGTACAAATGAATATATGTGTATGATGTTGCTCAGGTAAT 2985
Qy 2908 atgcagttcttctgaaggtgcttccatcacaga 2944
Db 2986 CTGAAGTACTCTGTAACCTGAAATCAACATATAGA 3022

RESULT 5
BOVPAS BOVPAS 3589 bp mRNA MAM 07-FEB-1999
LOCUS Bovine poly(ADP-ribose) synthetase (EC 2.4.2.30) mRNA.
DEFINITION D90073
ACCESSION D90073 GI:217581
VERSION DNA binding protein; helix-turn-helix; nuclear location signal;
KEYWORDS poly(ADP-ribose) synthetase; zinc-binding finger motif.
SOURCE Bovine thymus, cDNA to mRNA, clone pSO-7.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 3589)
AUTHORS Saito, I., Hatakeyama, K., Kido, T., Ohkubo, H., Nakanishi, S. and Ueda, K.
TITLE Cloning of a full-length cDNA encoding bovine thymus

poly(ADP-ribose) synthetase: evolutionally conserved segments and their potential functions
Gene (1990) In press
These data kindly submitted in computer readable form by: Isao Saito
Department of Clinical Science and Laboratory Medicine Faculty of Medicine, Kyoto University
Shogoinkawahara-cho 54, Sakyo-ku
Kyoto 606
Japan
Phone: 075-751-3467
Fax: 075-771-4792
Zinc-binding finger motifs are observed in seq. 21-51 and 128-165.
The sequences, 200-220 and 250-270, showed helix-turn-helix structure. Nucleotide-binding fold was found in seq. 890-903.

FEATURES

source Location/Qualifiers
1..3589 /organism="Bos taurus"
/db_xref="taxon:9913"
77..3127 /note="poly(ADP-ribose) synthetase"
/codon_start=1
/protein_id="BAAL4114.1"
/db_xref="GI:217582"
/translation="MAESSDKLYRVEYAKSGRACKKESIPKDSIRMAFVESPMP
DGKIPHWHLSCFWKVGFSIWPDPVEVEGFSSELWDQDTIKKAAETGGRTDVSQKQ
DGVSKTEKTLIDFGAGYAKSNRSTCKSMEXIDRGVRLSKVYYPDKPOLGMYDCW
YHPKCFVOKREELGPRPSPATHLMGFSVLTAEODETLKCOLPAIKGKRKRGDVG
IDEVYKKSKKEKDEIKLEKALKAQNDLIWNVDELAKKACSTLKKELLIFNKQEV
SESAILDRVAGMVFGLLPCEEGSGQLYFKGDAYCTGDTANTKCMVKTQTPNRK
ENWTKPREISYFKKTKIKODRIFPESPTPGAAPPSAASAPAAVHSGPPDKPL
SNMKILTLGKLSKNDKVAETKLEKGLTGTANKASLCISTKKEKLNKKMEVEKE
ANIRVSEDFLODIASATKSLQELLSTHLLSPWAEVPEVPEVAVGPKGGAAPSK
SKGPVKEGTNKSERKMLTKGGAADVPDGLSEHNAHVLEKGVFVATLGLVIVK
GTNSYKLOLLEDDRESYWIPIRSGVCTVIGSNKLEOMPSEKDAIEHFWKLYEKT
GNAAHSKNFTKHPKFEYLEIDYGQDEEAVKLTVPNGTSKLPKPVONLKMIFDVE
SNKAAVEYEIDLPKPLKNNANSVQAKVEMLNLDIEVAYSLLRGGSDSDSHLDSNRYT
LIPHPGFMKPPPLNNANSVQAKVEMLNLDIEVAYSLLRGGSDSDSHLDSNRYT
LKTIDIKVVDKDEAEIIRKYVKNTHATHNAYDLEVDVIFKIEREGESRYKPKQL
HNRLIWHGSRITNFAGILSOLRITAPPEAPVGYMGFYFADVMVSKSANYCHTSQ
GDP IGLILGEEAALGMVYELKHARHISKLPKGVHSGVGLTKTTPDPSASITVDGVEVP
LCTGISSGVNDTCLLYNEYIYVYDIAQVHLKYLKLEKFNKLSLW"
misc_feature 224..231 /note="nuclear location signal"
BASE COUNT 975 a 879 c 1010 g 725 t
ORIGIN

Query Match 8.3%; Score 245.8; DB 4; Length 3589;
Best Local Similarity 51.3%; Pred. No. 2.6e-52;
Matches 703; Conservative 14; Mismatches 614; Indels 39; Gaps 6;
Qy 1587 cnaggtgtgacaggtctactgtactcagatcattgaacaggtgatgggtctgagtg 1646
Db 1771 CAAGGACCAACTCTATTACAAGCTGCAGCTCCTGGAGGATGACAAAGAGAGAGCA 1830
Qy 1647 ctacgtattctgaagtgaggcgggttgaggatgagaaaatttggaggcgaactgga 1706
Db 1831 CTGGATATTTCAGGTCTCTGGGGCGGTGTGGCACGGTGATTGGTAGTAACAAACTGGAGCA 1890
Qy 1707 ggagatgtcaaaactgagcgaatcaaggaattcaaaagattattctctgagagactgg 1766
Db 1891 GATGCCATCAAGAGAGATGCCATTTGAGCAATTTATGAATTTATATGAAGAGAAAACCGG 1950
Qy 1767 aaactcatgggaagcttggaatgtaaaacaaattttcgaagcagcctgggagatttta 1826
Db 1951 AAAGCCTGGCACTCT-----AAAACTTTACGAAGCATCCCCAAAAAGTTCTTA 1998
Qy 1827 cccacttgatgttattgattgttgaagaagcaccac-----aacgggaagatcatcag 1980
Db 1999 CCCTCTGGAGATTGACTACGCCAGGATGAAGAGCGGTGAAGAAGTTGACAGTAACCC 2058
Qy 1881 tgaatgaaagtctctctctcctcaatgctagaactcatgaagatgcttttcaatgt 1940


```

Db 2059 TGGCAACAGTCCAGCTCCCAAGCAGCTGCAGACCTCATTAAGATGAATCTTTGATGT 2118
QY 1941 ggagacataagagctgctatgatgaatttgaataatgtcagaaatgctcttgg 2000
Db 2119 AGAGAGTATGAAGAAAGCCATGGTGGAGTATGAGATGACCTTCAAGAGATCCCTTGGG 2178
QY 2001 gaagctaaagcagaaatattgaaagatttgaagcattaaactkgrtacmrat 2060
Db 2179 GAAGCTGACCAAAAGCAGATCCAGCTGCATCTCCATCTCCTGATCTCTCAACCGCTT 2236
QY 2061 atttgaagacacccgctnatcaagcactgcttggtrgagaagctnaattgttgytsg 2120
Db 2237 -----GCACGTCTCCAGGCGAGCAGTGCATCTCACATCTCTGGATCTCTCAACCGCTT 2289
QY 2121 agmatssytttttactcttattcccttcttattcccttcttatttcaattatcagcgatgagat 2180
Db 2290 CTACACCTGATCCCGCAGGACTTCGGGATGAAGAAAGCCCGCTGCTGAACAAAGCAAA 2349
QY 2181 gatttcatataagcagaaatgcttgaagctctgcaggtattgaaattgcttcaaa 2240
Db 2350 CAGCGTGACGGCCAAAGTGGAAATGCTAGACAACTCTGCTGGATATTGAGTGCCTACAG 2409
QY 2241 gata-----gttggcttgcagcagctgagcttcttcttgcagctgataataatgaa 2294
Db 2410 TCTACTTAGGGTGGTCTGATGACAGCAGCAGAGGACCCATTGATGTCAACTATGAGAA 2469
QY 2295 acttactgtgacatacccccgtgctcagcagatgtagaattgaaattgagca 2354
Db 2470 GCTCAAACTGACATTAAGTTTGGACAAAGATTCCGAAGAGCGGAGATCATTAGGAA 2529
QY 2355 gtatctctcaacacatgctctactcaacagagctggtcgctgaaactgaggaagt 2414
Db 2530 GTATGTGAAGAACACTCACCGACCCACACAAACGGGTAGGACTTGAAGTCTGCACAT 2589
QY 2415 ttttctactgtgcagatgagaacttaataagtagtcaagataataataatctgca 2474
Db 2590 CTTCAAGATAGCGCGAAGGGAAAGCCAGCGTTTACAAAGCCGTTT---AAGCAGCTGCA 2646
QY 2475 taacaagatgctattatggcaaggttcaaggttgaagaaatttggtagaattttagtca 2534
Db 2647 TAACGGAGGCTGTGTGGACGGTCCAGGACCCACCAACTTCGCGGGCATCTGTGCCCA 2706
QY 2535 aggcctaagattgcactctgagcagctgttactggtctatgttcgcaagagcct 2594
Db 2707 GGGTCTCCGGATAGCCCACTGAGCAGCTGTGACGGGCTACATGTTTGGTAAGGGAAT 2766
QY 2595 ctacttgcagatctagtaagcagcagcagcagcagcagcagcagcagcagcagcagc 2654
Db 2767 CTATTTCCGGGACATGGTCTCCAAAGAGTGCCAACTACTGCCACAGCTCCCGAGGAGACC 2826
QY 2655 ttaggtttgagctgttcttctgaggttctttagagacatattgaactaaagaaagc 2714
Db 2827 AATAGGCTTGATCTGTGGGAGAGCTGCCCTTGGAAACATGATGATGAATTAAGACATGC 2886
QY 2715 cagctcattgacaaacctccaaagaggaagcatttcacaaagaggttaggcaaaacgt 2774
Db 2887 TCGACATATCAGCAAGTTTACCACAGGGCAGCAGCAGTGTCAAAGGTTTAGGCAAAAC- 2943
QY 2775 gccactgagtcagagtttgaagtgagagagagagagagagagagagagagagagagag 2834
Db 2944 TACCCCTGACCCATCGGCTAGTATTACTGTGGATGGTGTGGAGGTGCTCTCGGGACAGG 3003
QY 2835 ggtgccatcatcaattagagctctgaactcatgataagtagtacctcgtctcaacacac 2894
Db 3004 GATTTTCATCTGGTGTATGACACTGTCTGTTGTATTAACAGATACATCTCTACGACAT 3063
QY 2895 atcccaggtgaagatgcagttcttctgaaaggtgctgttccatcacaaaga 2944
Db 3064 TGCTCAGGTCCACCTGAAGTACCTGCTGAAGCTGAAGTTCAACTTTAAGA 3113

```

RESULT 6

```

XELPADPRP 3718 bp mRNA 04-FEB-1999
LOCUS Froq mRNA for poly(ADP-ribose) polymerase.
DEFINITION D14667
ACCESSION D14667.1 GI:287584
VERSION poly(ADP-ribose) polymerase.
KEYWORDS Xenopus laevis egg, oocytes, germ cell, cDNA to mRNA, clone
SOURCE XPAP5-9R.
ORGANISM Xenopus laevis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
Uchida, K.
Direct Submission
Submitted (15-MAR-1993) to the DDBJ/EMBL/GenBank databases.
Kazuhiko Uchida, Institute of Basic Medical Sciences, Department of
Biochemistry, Ten-noudai 1-1-1, Tsukuba, Ibaraki 305, Japan
(E-mail: kzuchida@emd.tsukuba.ac.jp, Tel.0298-53-3271,
Fax:0298-53-3039)
2 (bases 1 to 3718)
Uchida, K., Uchida, M., Hanai, S., Ozawa, Y., Arai, Y., Kushida, S. and
Miwa, M.
Isolation of the poly(ADP-ribose) polymerase-encoding cDNA from
Xenopus laevis: phylogenetic conservation of the functional domains
Gene 137 (2), 293-297 (1993)
94131300
3 (bases 1 to 3718)
Uchida, K., Uchida, M., Hanai, S., Ishikawa, K., Ozawa, Y., Ueno, N. and
Miwa, M.
Isolation of Poly(ADP-ribose) Polymerase cDNAs from Xenopus laevis
and Cherry Salmon using heterologous oligonucleotide consensus
sequences: gene conservation in amphibia, fish, and insect
Unpublished (1993)
Submitted (15-MAR-1993) to DDBJ by:
Kazuhiko Uchida
Department of Biochemistry
Institute of Basic Medical Sciences
University of Tsukuba
1-1-1 Ten-noudai
Tsukuba, Ibaraki 305
Japan
Phone: 0298-53-3271
Email: kzuchida@emd.tsukuba.ac.jp
Fax: 0298-53-3039.
FEATURES
Location/Qualifiers
1..3718
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="Germ cell"
/dev_stage="Egg"
/tissue_type="oocytes"
1..3718
/standard_name="Xenopus poly(ADP-ribose) polymerase"
BASE COUNT 1105 a 798 c 954 g 861 t
ORIGIN
Query Match 8.1%; Score 239.2; DB 5; Length 3718;
Best Local Similarity 51.8%; Pred. No. 1.3e-50;
Matches 701; Conservative 12; Mismatches 602; Indels 39; Gaps 7;
QY 1602 ctactgtactccagatcatgaacagagatgaggtctgagctgctacgtatttcgtaa 1661
Db 1787 CTTATTATAAGCTCAACTTATAGACGACGACAGATTCAGGTACTGGGTGTTCGATC 1846
QY 1662 gtggagcaggttggagtgagaaattgaggcgcaaaactggagagatgtcctcaaaac 1721
Db 1847 CTGGGTGCGGTGGGACACAGTATAGCGCATTAAGCTAGAGAGATGTCTTCCAAGGA 1906
QY 1722 tgaggcaatcaaggaattcaaaagattatttcttgagagagactggaaactcatgggaac 1781
Db 1907 AGATGCCATTGAACATTTCTCAACCTGTATCAGGACAAACCTGGCAATGCATGGCCTC 1966

```

QY 1782 ttgggaatgtataaaccaatttttcggaagcagctggtgagatgttttaccactgtatgtga 1841
 Db 1967 GC-----CCAAATTCACCAATACCCCAAAATTCATCCCTCGGAATAGA 2014
 QY 1842 ttatggtgttaagaagaccacaaacggaa-----agatatcagtgaataaagaatttc 1895
 Db 2015 CTATGACAGGAGAGAGATGTGTGAAGAAGCTCTCAGTGGGGCGCGGCACAAATCCAA 2074
 QY 1896 tctgtctcctcaatgtcagactcatgaagatgcttttcaatgtgagacatatagagc 1955
 Db 2075 GCTCGGTAGCTGTCCAGCAACTGATCAATAATTTTGTATGTGAGAGCATGAAAAA 2134
 QY 1956 tgctatgatggaatttgaawtaatatgttcagaaatgctctctgtggaagctaaagcmaggr 2015
 Db 2135 GGCCATGTGTGAGTTGAGATGATCTCCAGAGATGCTTTGGGAAACTCAGCAACGC 2194
 QY 2016 aaatattgagaagattgattgaagcattactkrgrtaacmratattattgaagcacccg 2075
 Db 2195 GCAGATTGAGAGTGGCTATCTTATTCTTAGCAAGTGCAGCAGCGCGTGTCTG---AATC 2251
 QY 2076 ctnatcaagcactgctgtttrgagaaagctnaattgttgytgagcmatssytttttc 2135
 Db 2252 CTTAAGCGAAGCCCATTTACTGATCTCTCAATCAGATTTTACACACTACCTCATGA 2311
 QY 2136 actctatoccttctattcctcctcattatattacggggatgaggatgatttctatcaaa 2195
 Db 2312 TTTTGAATGAAGAAGCCCGCTGCTAAATAACCTAGATGACAT-----ACAGGCTAA 2365
 QY 2196 ggcgaataatgtgaagctgcaggaatattgaattgtctcaa-----agatagttg 2249
 Db 2366 AGTGAGATGCTGGGATATCTGCTGACATTTGCTGACATTTGCTTACAGCCTGTTGAGAGTGG 2425
 QY 2250 cttegatgcagcagctgatgaatctcttgatgataaaatatatgaactcacttgacat 2309
 Db 2426 CCGCCGATGATGCTGAAAAGGATCCCATTTGATGTGAATATGAAAGATTAGACTGACAT 2485
 QY 2310 caccocctgctcagcatagtagaataacagtttaattgagcagtatctcctcaaac 2369
 Db 2486 TAAGTTTCTGCTAAGATTCAGAGAATCCAGAAATATATGCGATTATGTCAGAAACAC 2545
 QY 2370 acatgctcactcacaagagctgctgctggaactggaagagtttttcaattgactg 2429
 Db 2546 GCAGCTGATAGCCACATGATGATCTGAGGTCTCTGAGATATCAAAATCGACCG 2605
 QY 2430 agatggagaacttaataagctactcaagataaaaataatctgcatacaagatgctatt 2489
 Db 2606 TCAAGGTGAATATCAGCGGTATAAACCATTTAAA---CAGCTACACAAACCGCGCTGCT 2662
 QY 2490 atgacaggttcaaggttgacgaattttgtgggaattcttagtgaagggttaagaattgc 2549
 Db 2663 TTGGCAGCGCTCCCGCACACAGAAATTTTGAGGAATATTCTCTCAGCGGTCTCCGAATTCG 2722
 QY 2550 acctcctgagggcactgttactgctatatgttcgcaaaagcctcacttttcagatct 2609
 Db 2723 TCCGCCAGAGCTCTGTTACCGGGTATATGTTTGGCAAGAGGTATCTATTTTCCGACAT 2782
 QY 2610 agtaagcaagagcacaactactgttatgttgataggaaataactcctgtaggtttgatgct 2669
 Db 2783 GGTATCAAAAGTGCAAACTACTGTCTCAATGCTCGGGTAGCCCATAGGGGTGATCTT 2842
 QY 2670 tctttctgaaggttctttaagagacatgtatgaactaaagaaagccacgtccatgacaa 2729
 Db 2843 ACTGGGGAGGTGGCCCTCGGAACAATGATGAGCTGAAGAGCAGCATCAAAATTCAAA 2902
 QY 2730 acctccaaggggaagcattccgcaagaggttagcacaacccgtccactggagtcaga 2789
 Db 2903 ACTTCCCAAGGGCAACACAGTGTAAAGGTTTGGGTAGACTGCACAGATCCCTCAGC 2962
 QY 2790 gtttgaagtgagggatgatctgtagtccctgcggcaagccggtgccatcatcaat 2849
 Db 2963 TACTGTACA---GCTGGATGGAGTGTATGTTCTCTCGGGAAGGAAGTCTCTGCAAAATAT 3019

QY 2850 taggagctctgaactcatgtacatgagtgatcatcgtctacacacatccccagggtgaagat 2909
 Db 3020 CAGTGACACCAGCCTGTGTATAATGAATATATTGTGTATGATATTGCTCAGTGCAACCT 3079
 QY 2910 gcagttctctgtagaggtggtggttttccatcacaa 2943
 Db 3080 GAAGTACCTGCTGAAGCTCAAGTTCACTACTACAAG 3113

RESULT 7
 XLPARPG
 LOCUS
 DEFINITION XLPARPG 3617 bp mRNA VRT 18-JAN-1993
 ACCESSION X.laevis PARP gene encoding poly(ADP-ribose) polymerase.
 VERSION Z12139.1 GI:64967
 KEYWORDS poly(ADP-ribose) polymerase.
 SOURCE African clawed frog.
 ORGANISM
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 3617)
 Saulier-le Dreaan, B.M.
 Thesis (1992) Lab. de Biol. et Genet. du Developpement, Universite
 de Rennes I. URA CNRS 256
 2 (bases 1 to 3617)
 Saulier-le Dreaan, B.M.
 Direct Submission
 Submitted (15-MAY-1992) Saulier-le Dreaan B.M., Lab. de Biol. et
 Genet. du Developpement, Universite de Rennes I. URA CNRS 256,
 Campus de Beaulieu, Av. du Gal Leclerc, Rennes-cedex, FRANCE, 35042

FEATURES
 Location/Qualifiers
 1..3617
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /germline
 /sex="Female"
 /dev_stage="Adult"
 /tissue_type="Ovary"
 /cell_type="Oocyte"
 /clone_lib="oligo dT primed lambda gt10 library of D.
 MELTON"
 /clone="clone parp 923"
 1..2997
 /gene="PARP"
 1..2997
 /gene="PARP"
 /EC_number="2.4.2.30"
 /function="poly ADP-ribosylation, DNA repair"
 /note="Protein sequence is in conflict with the conceptual
 translation: poly(ADP-ribose) polymerase"
 /citation=[1]
 /codon_start=1
 /product="NAD(+) ADP-ribosyltransferase"
 /protein_id="CAA78126.1"
 /db_xref="GI:1334661"
 /db_xref="SWISS-PROT:P31669"
 /translation="AKSRASCKKCGDNIAKESLGIALMVQSPMFDKGVPHWHYSCF
 WKRAVLSQDIYITELRWEDQEMIKKAIETGGAAGAGDSDGKGKGMTLNDFAAE
 YAKNSACKGCEQIKIQKQIRSKSDVERPOLQIDRMVHPDCEVSSREELDFLP
 YASQALQKGTILSAEDSLKLLPAKNEGRKRADEVGHSAAATKKIKKKEKES
 KLEKLEKELIWHIKDELKVCSTNDLKEHLIANQKQVSGETNIVDRVSDGMAG
 ALLPECSQGVFKGDAYCTGLSAATKCAKTOPTNRKDMVTPEFHEIDYLLKF
 KFKRDRAFPCCAPTPLSPAAPEPKPTVEETPECKPLTNTKVLIGLKLKSNKDE
 KTLLEGCGVAGSAHAKNICISTNKEVKMSKMEVKAANRVVSDDFLKEVESK
 SVQELLQFGLSSWGAIEQEAQVPTKQSPGVAGKSSGKSKKSKKMKL
 TVGKGAIDPDSLEDSCHVLETGKTSATLGLVDITRTGNSYIKLLEHRSRY
 WFRSGRVGTIVGSKLEEMSKEDAIEHFLNODKTGNWHSNFTYKPKFPL
 EIDYQIEDVVKLSVAGTKSLAKPVQELIKLIPVESMKKAMVEFEDLQKPLG
 KLSKROIQSAVSILSQVQAVSLSLSEARLLDLSNOFTLLPHDFGKMKPPLNLEY
 IQKVMLDNLLDIEAVSLRGGADGCKDIDVYEKIKTDIKVAKDSESRITIC
 DYKNTHADTHNADLEVLKIFKIDREGEYQRYKPKQLNRLQMLHSGSRRTFAGIL
 SQGLRIAPPEAPVTGYMFGKIYFADMVSKSANYCHAMPGPSIGLILGLVALGNMHE."

gene

CDS

LKASQTKLPKHKSHVKGRLTRAPDPSATVOLGDVDPVPLGKTSANISDTSLLYNEY
IYDIAQVNLKYLKLLKFNKYGMMW

3'UTR
polyA_signal
polyA_site
BASE COUNT 1131 a 757 c 910 g 819 t
ORIGIN

Query Match 8.1%; Score 238.2; DB 5; Length 3617;
Best Local Similarity 51.7%; Pred. No. 2.4e-50;
Matches 700; Conservative 12; Mismatches 602; Indels 39; Gaps 7;
QY 1602 ctactatgtactccagatcattgaacagagatgatgggtctgaagtcactcgtatcttcgtaa 1661
DB 1653 CTATTATAAGCTGCAACTTATAGACGACAGAGATTCACAGTACTGGGTGTCCGATC 1712
QY 1662 gtggagacgggttgagtgagaaatggagggcaaaactggaggagatgcataaac 1721
DB 1713 CTGGGCTGGGTGGGACAGTAAAGGCAGTAAAGAGTAAAGAGATGCTTCCAAAGGA 1772
QY 1722 tgaggcaatcaaggaaattcaaatgattttcttgagagagactgaaactcactggaagc 1781
DB 1773 AGATGCCATTGAACATTTCTCAACCTGATCAGGACAAACTGGCAATGCAATGGCACTC 1832
QY 1782 ttgggaatgaaacccaatttttggaagcagcctggagagattttaccacttgatgtga 1841
DB 1833 GC-----CCAAATTCACCAATACCCCAAAATTTCTATCCCTCGAAATAGA 1880
QY 1842 ttatgtgttaagaagacacca-----aaacggaagatatcagtgaaatgaaagttc 1895
DB 1881 CTATGACAGGAGAGAGATGTGTGGAAGAACTCTCAGTGGGGCGCGCACAAATCCAA 1940
QY 1896 tcttgcctcaattgtctgagaactcatgaagatgcttttcaatgttgagacatatagac 1955
DB 1941 GCTCGCTAAGCCTGCTCCAGGAAGTATCAAAATTAATTTTGTGATGGAGAGCATGAAAA 2000
QY 1956 tgcatagtgaatttgaaatwaatgtcagaaatgctcttggaagcctaagcmaggr 2015
DB 2001 GGCATGTGGTGGATTTGAGATCCATCTCCAGAGATGCTTTGGGGAACCTCAGCAAGC 2060
QY 2016 aaattattgaaagatttgaacattaaactkrtrtaacmrattttatttgaagacacg 2075
DB 2061 GCAGATTACAGTGCCTGCTCAATTTCTAGCCAAAGTACGACGCGGTGCTG---AATC 2117
QY 2076 ctnatcaagcactgctgtttrgagaagcnaattgttgytsagmatssyttttc 2135
DB 2118 CTTAAGCGAAGCCGGATTACTGGATCTCTCAATFCAGTTTTACACACTCATCTCATGA 2177
QY 2136 actcttacccttcttcttccatctattatcacgggatgagatgatttcattcaaa 2195
DB 2178 TTTTGGATGAAGAAGCCCGCTGCTGCTAAATAACCTAGATGACAT-----ACAGGCTAA 2231
QY 2196 ggcgaataatgctgaagcctctgcaggtatgaattgcttcaa-----agatagttg 2249
DB 2232 AGTCAGATGCTGATTAATCTGCTGACATTTGAAGTTGCTTACAGCCTGTTGAGAGTGG 2291
QY 2250 cttcgatagcacagtgatgaactctctgtatataatataatgaactcactgtgacat 2309
DB 2292 CGCCGATGATGGTGAAGAAGATCCATTGATGTGAATATGAAAGATTAAGACTGACAT 2351
QY 2310 caccocctggctcacagtagtgaagattacaagtttaattgagcagtatctctcacaac 2369
DB 2352 TAAGGTTTGTCTAAGATTTCAGAGAATCCAGAATCATATGCGATTATGTCAAGAACAC 2411
QY 2370 acatgctcctactcacaagactggtcgctggaactggagaaatttttttcaactgacg 2429
DB 2412 GCACCCAGATCGCACAATGATGATCTTGAGGTCTCCAGATATTCAAATCGACCG 2471
QY 2430 agatggagaacttaataagtaactcaagataataaaataatctgcatacaagatgctatt 2489
DB 2472 TGAAGGTGAATATCAGCCGGTATAAACCACTTTAA---CAGCTACACAAACCCGACGCT 2528

QY 2490 atggcacggttcaaggttgacgaatttttgggaattcttagtcaaggcgtcaagaattgc 2549
DB 2529 TTGGCAGCGCTCCCGCACCAAAATTTTGCAGGAATATTGCTCAGGGTCTCCGAATTC 2588
QY 2550 acctctgagcaccctgttactggtctatatttcgcaaaaggcctcacttctcagatct 2609
DB 2589 TCCGCCAGAAAGCTCTCTTACCGGGTATATGTTGGCAAGGATATCTATTTTCCGCACAT 2648
QY 2610 agtaagcaagagcgacacaaactactgttattggtataggaataaactctgtgattgtgct 2669
DB 2649 GGTATCCAAAAGTGCAAACTACTGTCTCATGCAATGCCGGTAGCCCATAGGGCTGATCTT 2708
QY 2670 tctttctgaggttcttttagagacatgtatgaactaaagaagcaacgctccatggacaa 2729
DB 2709 ACTGGGGAGGTGGCCCTCGCAAAACATGATGCTGAGCTGAAAGCAGCATCACAAATACAAA 2768
QY 2730 acctccaaagaggaagcatttcgacaaaggttaggcaaacctgacactggagtcaga 2789
DB 2769 ACTTCCCAAGGGCAACACACAGTGTAAAGGGTTTGGGTAGAACTGCACAGATCCCTCAGC 2828
QY 2790 gtttgtgaagtggaggatgatgctgtagttccctcgccgaagccggtgccaatcaat 2849
DB 2829 TACTGTACA---GCTGGATGGAGTGGATGTTCTCTCGGAAAGGAACTTCTTGCAAAATAT 2885
QY 2850 taggaagctcctgaactcatgtacaaatgagtagatcgtctacaacacatcccaggtgaagat 2909
DB 2886 CAGTGACACACAGCCTGTTGTATATGAATATTTGTGTATGATATTGCTCAGGTCAACCT 2945
QY 2910 gcagttctgctgaaggtgcgtttccatcacaa 2942
DB 2946 GAAGTACCTGCTGAAGCTCAAGTTCACACTACAA 2978

RESULT 8
HUMPOLP HUMPOLP 3640 bp mRNA PRI 08-JAN-1995
LOCUS Human poly(ADP-ribose) polymerase mRNA, complete cds.
DEFINITION M18112
ACCESSION M18112.1 GI:190166
VERSION polymerase.
KEYWORDS Human SV40 transformed fibroblast, cDNA to mRNA, clone pPAP.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3640)
AUTHORS Uchida,K., Morita,T., Sato,T., Ogura,T., Yamashita,R., Noguchi,S.,
Suzuki,H., Nynuya,H., Miwa,M. and Sugimura,T.
TITLE Nucleotide sequence of a full-length cDNA for human fibroblast
poly(ADP-ribose) polymerase
JOURNAL Biochem. Biophys. Res. Commun. 148 (2), 617-622 (1987)
MEDLINE 88076933
COMMENT Draft entry and computer readable sequence for [1] kindly provided
by K.Uchida, 02-MAR-1988.
FEATURES
Location/Qualifiers
1..3640
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="lq41-q42"
<1..3640
/note="POLP mRNA"
140..3184
/gene="PPOL"
140..3184
/gene="PPOL"
/note="poly(ADP-ribose) polymerase"
/codon_start=1
/db_xref="GDB:G00-119-508"
/protein_id="AAA60137.1"
/db_xref="GI:190167"
/translation="MAESSDKLYRVEYAKSGRASCCKSESIPKDSLRMAIMVQSPMF
DKQVPHWYHFCFVKVGHSHIRHPDVQVDFSELRWDDQVKVKTAEAGCVTKGQDGI
GSKAEKTLGDFAAEYAKSNRSTCKGCMKLEKGOVRLSKMVDPEKPOLGMDIRWYHP
GCFVKNREELGFRPEYSQKGLFSLATEDKEALKKQLPGVKSGKRKGEVDGVDGE

BASE COUNT 999 a 833 c 1008 g 800 t
ORIGIN Chromosome lp11-qter.

Query Match	7.6%;	Score 224.6;	DB 9;	Length 3640;
Best Local Similarity	51.1%;	Pred. No. 7.8e-47;		
Matches 693;	Conservative 11;	Mismatches 612;	Indels 39;	Gaps 7;
QY 1602	ctactatgtactccagatcattgaacagagatgagggtctgagtgctacgtattcgttaa	1661		
DB 1843	CTACTACAAAGCTGCGAGCTTCTGGAGGACGACAAAGGAAACAGGATATTGGATATTCAAGTCC	1902		
QY 1662	gtggggacgggtgtggagtgagaaattggagggccaaataactggaggagatgtcaaaaac	1721		
DB 1903	CTGGGGCCGTGTGGGTACGGTGATCGGTAGCAACAACTGGCAACAGATGCCGTCCACGGA	1962		
QY 1722	tgaggcaatcaaggaattcaaaagattattcttgagaagactggaactcatgtgggaagc	1781		
DB 1963	GGATGCCATTGAGCACTTTCATGAATTTATATGAAGAAAAACCGGAACGGCTTGGCACTC	2022		
QY 1782	ttgggaaatgtaaaaccatttccggaagcagccctgggagattttaccactcttgatgtga	1841		
DB 2023	C-----AAAAATTTACGAGATATCCCAAAAAGTTCTACCCCTCGGAGATTGA	2070		
QY 1842	ttatgtgtttaagaaagcacc-----aaacggaaagatatcagttgaaatgaaaagttc	1895		
DB 2071	CTATGGCCAGGATGAAGGCGAGTGAAGAAGCTGACAGTAAATCTTGGCACCAGTCCAA	2130		
QY 1896	ctctgctcctaattgctagaactcatgaagatgctcttccaatgtggagacatatagagc	1955		
DB 2131	GCTCCCNAGCCAGTTCAGGACCTTCATCAAGATGATCTTTGATGTGGAATGTAAGAA	2190		
QY 1956	tgctatgatgaaattgaaawtaatatgtcagaatacctctctgggaagctaagcmaggr	2015		
DB 2191	AGCCATGGTGAGTATGAGATCGACCTTCAGAAGATGCCCTTGGGNAAGCTGAGCAAAAG	2250		
QY 2016	aaatattgaaaggaatttgaagcattaactckrgrtaomgratttatttgaaggaacacq	2075		
DB 2251	GCAGATCCAGCGCGCATACTCCATCTCTAGTGAAGTCCAGCAGCGGGT-----GTCTCAG	2305		
QY 2076	ctnatcaagcactggctgtttrgagaaagctnaattgtgtgtsgagcmatsytttttc	2135		
DB 2306	GGCAGCAGCCACTCTCAGATCTCTGGATCTCTCAAAATCGGTTTACACCCGTGATCCCCCAC	2365		
QY 2136	actettatcccttctatctcatctcatattatcagggatgagatgattcatattcaaa	2195		
DB 2366	GACTTTTGGGATGAAGAGGCTCCGCCCTCTGAACAAATGCAGA-----CAGTGTGCAGGCCAA	2421		
QY 2196	ggcgaaaatcttgaagcctcgagatattgaaattgcttcaa-----agatagttgg	2249		
DB 2422	GGTGGAAATGCTTGACAACTGCTGTGGACATPCGAGGTGGGCTACAGTCTGCTCAGGGGAGG	2481		
QY 2250	cttcgatgcagcagtgtgaattctcttgatgataaataatgaaacttcaactgtgaacat	2309		
DB 2482	GTCTGATGATAGCAGCAAGGATCCCATCGATGTCAACTATGAGAGCTCAAAACTGACAT	2541		
QY 2310	caccccgctggctcacgatagtgaagattacaagttaatgtgacgatctctctcaaac	2369		
DB 2542	TAAAGTGGTTGACAGAGATCTTGAGAGGCCGAGATCATCAGGAAGATGTTTAAAGAACAC	2601		

QY	2370	acatgctcctactcacaaaggactggtcgctggaaactggagggaagttttttcacttgatcg	2429
Db	2602	TCATCAACCACACACAAATGCGTATGACTTGGGAAGTCATCGATATCTTTAAGATAGACGC	2661
QY	2430	agatgagaaacttaataagttactcaagatatataaaataatctgcataacaagatgctatt	2489
Db	2662	TGAAGGCCGAATGCCAGCGTTTACAGGCCCTTT---AAGCAGCTTCATAACCGAAGATTGCT	2718
QY	2490	atggcacggttccaaggttgacgaattttgagaaattcttagtcaagggtcaagaattgc	2549
Db	2719	GTGGCAGGGTCCAGGACCAACACTTTGCTGGGATCCTTGCCCCAGGGCTCTCGGATAGC	2778
QY	2550	acctctgaggcaactgttactggctatatgttcgcgaaagcctctactttgcagatct	2609
Db	2779	CCCGCCTGAAGCGCCGTGACAGGCTACATGTTTGGTAAAGGGATCTATTTCGCTGACAT	2838
QY	2610	agtaagcaagagcgacaataactgtttatgtgtagaggaataatcctgttagtttgatgct	2669
Db	2839	GGTCTCCAAGAGTGCCAACTACTTGCATACGTCTCAGGGAGACCAATAGGCTTAAATCCT	2898
QY	2670	tctttctgaggttgctttaggagacatgtatgaactaaagaaagccacgctccatggacaa	2729
Db	2899	GTTGGGAGAGTTGCCCTTGGAAACATGTATGAACCTGAAGCACGCTTCACATATCAGCAA	2958
QY	2730	acctcaagagggaagcattcgcacaaaggattaggcaaaaccgtgcactggagtcaaga	2789
Db	2959	GTTACCCAAGGGCAAGCACAGTGTCAAGGTTTGGGCAAAACTACCCCTGTATCCTTCAG-	3017
QY	2790	gtttgtgaagtggaggatgatgtcgtagttccctcgcgcaagccgctgcaatcaaat	2849
Db	3018	--CTAACATTAGTCTGGATGGTGTAGAGCTTCCCTTTGGGACCGGGATTTCATCTGGTGT	3075
QY	2850	tagtagctctgaactcatgacaatgagtacatcgctctcaacacatcaccagtggaagt	2909
Db	3076	GAATGACACCTCTCTACTATATTAACAGGTACATTGTCTATGATATTCCTCAGGTAANTCT	3135
QY	2910	gcagttcttgtaagggtgcgtttccatcaacaaga	2944
Db	3136	GAAGTATCTGCTGAACCTGAAATCAATTTTAAAG	3170

```

RESULT      9
LOCUS       A52134
DEFINITION  A52134 3792 bp DNA
ACCESSION   A52134
VERSION     A52134.1 GI:2304739
KEYWORDS    .
SOURCE      unidentified,
            unidentified
            unclassified.
ORGANISM    1 (bases 1 to 3792)
REFERENCE   1 (bases 1 to 3792)
AUTHORS     Buerkle,A., Zur.H.H. and Kuepper,J.
TITLE        VECTORS AND VIRUSES FOR USE IN GENE THERAPY
JOURNAL      Patent: WO 9618737-A 1 20-JUN-1996;
            DEUTSCHES KREBSFORSCH (DE)
FEATURES    Location/Qualifiers
             1..3792
             /organism="unidentified"
             /db_xref="taxon:32644"
             96...3137
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="CAA03385.1"
             /db_xref="GI:2304740"
            CDS

```

/dev.x86_64: "GI.2.304740"
 /translation: "MAESSDKILRYEYASERASCKSCSEIPKDSLRMAIIVQSPMF
 DGKYPHWHSFCAWYKHSIRHPDEVGDSLSRWDDQSKVKYAEAGVTQGGQDGI
 GSAEKTLTDFEYKAYASNSRSTCKMEKIEKQVLSKMKVDEAKQGLMDIRWHP
 VPKVKEELGFPEYASQGLSILATEDEKALKQOLPGVSEKGRKGDVQVDE
 GAKKSKKDKDQSLKLEKRLQNLINWIKDEALKQVSKLDELKILFNKQVPSGE
 TIRLADVADGMEFGALLCEESGQGLFYSDAYVTGDTVAATCMVKYKTOPNKKW
 TPKEFRFISYLLKLVKVKQDRIIPFETSAASVAVHPPTSAAPAAVNSADPKLSKM

KILTLGLKLSRNKDEVKAMIEKLGKLTGTANKASLICIISTKKEVERKMKMBEVEKREANI
 RVSEDFQDNVSAKSLQELFLAHLISPLGAEVKAEPVEVAPRCKSGAALSRRKSG
 OYKESICLNSEKRLITLGGAAVDPDSLEHSAHLEKGGKVFSAITGLIVKGTN
 SYKQLLEDDEKRENYWIFRSWGRVGTIGSNKLEOMPSKEDAIEHFMKLYEEXTGNA
 WSKNFYTPKRPFLPDIYDGEAEVRLITVPNGTKSLKRPVODLILKMLFEDVESMK
 KAMVETIDLOKMPLEIGKSKRQIOAYISILSEVOAASOGSSDILDLNKNFYLLIP
 HDGKMKPPLNNAADSVQAKBMLNDLIDIEVAYLSLRGSDSDSKDPIDVNYEKLKT
 DIKVVDRDSEEEAIRKVKYKNTHTHAYDLIEDIFKIEREGECORYKPFKQLHNR
 RLLWHSRTNLFAGILSOLRTAPPEAVPTVMFGKIGYFADVMSKSNYCHTSGODP
 IGLILLGEVALNMVELKHASHIKLPKSHVKGIGKTPDPDSANISLDGVDVPLGT
 GLISSVNDTSLINIEYIVDIAQVNLKYLKLFNFKISLW"
 BASE COUNT 1048 a 846 c 1035 g 863 t
 ORIGIN

Query Match 7.6%; Score 224.6; DB 6; Length 3792;
 Best Local Similarity 51.1%; Pred. No. 7.8e-47;
 Matches 693; Conservative 11; Mismatches 612; Indels 39; Gaps 7;

QY 1602 ctactgtactccagatcattgaacaggatgatgggtctgagctgactgatttcgttaa 1661
 DB 1796 CTACTCAACGCTGCAGCTTCTGAGGACGACACAGGAAACAGGTATTGGATATTCAAGGTC 1855
 QY 1662 gtgggacgggttggagtgagaaaaattggaggggcaaaactggagagatgtcaaaaac 1721
 DB 1856 CTGGGCGCGTGTGGGTACGGTGTATCGGTAGCAACAACACTGGAACAGATCCGCTCCAGGA 1915
 QY 1722 tgaggaatcaagaatacaagattattcttgagagactggaaactcatgaggagac 1781
 DB 1916 GGATGCCATTGAGCACTTCATGAAATATATAGAGAAACCCGGAACGCTTGGCACTC 1975
 QY 1782 ttgggaatgtaaaaccaaattttggagagcggcggtgagattttaccacttgattga 1841
 DB 1976 C-----AAAATTTCAGGAGTATCCCAAAAAGTTCTACCCCTGGAGATTGA 2023
 QY 1842 ttatggtgttaaaagacacca-----aaacggaaagatatcagtgaaatgaaagttc 1895
 DB 2024 CTATGGCCAGGATGAAGAGGACGTGAAGAAGCTGACAGTAAATCCTGGCCACCAGTCCAA 2083
 QY 1896 tctgtccctcaattgtagaactcatgaagatgcttttcaatgtggagacatatagac 1955
 DB 2084 GTCCTCCCAAGCCAGTTCAGACCTCATCAAGATGATCTTTGTATGTGGAAGTATGAAGAA 2143
 QY 1956 tgcattgatgaatttgaataatattgtcagaaatgcctcttgggaagcgaagcaggr 2015
 DB 2144 AGCCATGTGGAGTATGAGATCGACCTTCAGAGATGCCCTTGGGAAGCTGAGCAAAAG 2203
 QY 2016 aaattatgagagagatttgaagcattaaactkgrtacmrratttatttgaagagacacg 2075
 DB 2204 GCAGATCCAGGCGCATACTCCATCTCTCAGTGGTCCAGGCTCCAGAGCGCGT-----GTCTCAG 2258
 QY 2076 ctatcaagcactggctgtttrgaagaaagctnaattgttgytsgagacmatssyttttc 2135
 DB 2259 GGCAGACGACCTCTCAGATCCTGGATCTCTCAATCGCTTTTACACCTTGATCCCCCACC 2318
 QY 2136 actctatccctcttattcatctctatattacgggagatgagatgatttcattcaaaa 2195
 DB 2319 GACTTTGGGATGAAGAAGCGCTCCCTCTGAACAAATCGA-----CAGTGTGAGGCGCAA 2374
 QY 2196 ggcgaataatgcttgaagctctgcaggaatttgaattgttcaaa-----agatagttgg 2249
 DB 2375 GGTGGAATGCTTGACACCTCTGGACATCGAGGTGGCTTACAGTCTGCTCAGGGGAGG 2434
 QY 2250 ctctgatagcagagtgatgaattctcttgatgataaataatgataaacttctcactgtgacat 2309
 DB 2435 GTCTGATGATAGACAGCAAGATCCATCGATGTCACTATGAGAGAGCTCAAACTGCAT 2494
 QY 2310 caccocctggctcagcagtagtgaagattacaagtttaattgacagtatctctcctaacac 2369
 DB 2495 TAAGGTGGTTGACAGAGATTCTGAAGAAGCGGAGATCATCAGGAAGTATGTTAAGAACAC 2554
 QY 2370 acatgtcctctactcaagaagactggtgcgtggaactggaagagttttttcacttgatcg 2429

DB 2555 TCATGCAACCAACACACAAATGCGTATGACTTGGAAGTCAATCGATATCTTTAAGATAGAGCG 2614
 QY 2430 agatgagaaacttaataagactactcaagataaaaaataatctgcatacaagaatgctatt 2489
 DB 2615 TCAAGGCCAATGCCAGCGTTACAAGCCCTTT--AAGCAGCTTCATAACCGAAGATTGCT 2671
 QY 2490 atggcacggtttaagggttgagcaaatatttttgggaattcttagtcaagggtctaaagaattgc 2549
 DB 2672 GTGGCAGCGGTCCAGGACCAACAACTTTGCTGGGATCTCTGCCAGGGCTCTTCGATAGC 2731
 QY 2550 acctctgaagcactgttactggtctatatgttcggcaaaagcctctacttgcagatct 2609
 DB 2732 CCCGCTGAAGCGGCCGCTGACAGGCTACATGTTTGGTAAAGGATCTATTTTCGCTGACAT 2791
 QY 2610 agtaagcagagcgcacaaactgttatgtgtaggagaaataatctctgtagtctgtagtct 2669
 DB 2792 GGTCTCCAGAGTGCACAACTACTGCATACGTCCTCAGGAGACCCAAATAGGCTTAATCCT 2851
 QY 2670 tcttctgaggtgtgttaggagacatgtatgaactaaagaagcagcagtcacatggacaa 2729
 DB 2852 GTTGGGAGAAGTTGCCCTTGGAAACATGTATGAACATGAAGCAGCGCTTCACATATCAGCAA 2911
 QY 2730 acctccaagaggaagcattcgaccaaggattagcacaacccgtgacctggagtcaga 2789
 DB 2912 GTTACCAGGCAAGCAGACAGTGTCAAGGTTTGGCAAAACTACCCCTGATCTTCAG- 2970
 QY 2790 gttgtgaagtggagggatgctgtagtctccctgcggcaagccggtgacctgcatcaat 2849
 DB 2971 --CTAACATTAGTCTGGATGGTGTAGACGTTCTCTTGGACCGGGATTTCTATCTGGTGT 3028
 QY 2850 taggagctctgacatcagtcacaaatgagtcacatcgctctacaacacatcccagggtgaagat 2909
 DB 3029 GAATGACACCTCTCTACTATATAACGAGTACATTCCTGATGATATTGCTCAGGATAAATCT 3088
 QY 2910 gcagttcttctggaaggcgcttccatcacacaaga 2944
 DB 3089 GAAGTATCTGCGTGAACACTGAATTCATTTTAAAGA 3123

RESULT 10
 HUMRISDAD
 LOCUS Human poly(ADP-ribose) synthetase mRNA, complete cds.
 DEFINITION J03473
 ACCESSION J03473.1 GI:337423
 VERSION poly(ADP-ribose) synthetase.
 KEYWORDS Human placenta, cDNA to mRNA (library of H.Okayama), clones
 SOURCE PPARS[1,11,21,32,32,41,-F].
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3795)
 AUTHORS Kurosaki,T., Ushiro,H., Mitsuuchi,Y., Suzuki,S., Matsuda,M.,
 Matsuda,Y., Katunuma,N., Kangawa,K., Matsuo,H., Hirose,T.,
 Miyama,S. and Shizuta,Y.
 TITLE Primary structure of human poly(ADP-ribose) synthetase as deduced
 from cDNA sequence
 JOURNAL J. Biol. Chem. 262 (33), 15990-15997 (1987)
 MEDLINE 88058958
 COMMENT Draft entry and printed copy of sequence for [1] kindly provided by
 Y.Shizuta, 23-NOV-1987.
 FEATURES
 source
 1..3795
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="1q41-q42"
 <1..3795
 /note="poly(ADP-ribose) synthetase mRNA"
 96..3140
 /gene="PPOL"
 96..3140
 /gene="PPOL"
 /note="poly(ADP-ribose) synthetase"

```

/codon_start=1
/db_xref=GDB:G00-119-508"
/protein_id="AAB59447.1"
/db_xref="GI:337424"
/translating="MAESDPLKLYVEAYKASERASCKKSESIPKDSILRMALIMVQSPMF
DGVPHWYHFCSPFWKYSIRHPDVEDGSEGLVRDQOQVKYKTAABAGVGTGKGQGI
GSAEKTLDGFAAEYAKNSHTRHPDVEDGSEGLVRDQOQVKYKTAABAGVGTGKGQGI
GCTVKNEELGPRPEVSASQLGFSLLATEDKEALKQLPGVSEGRKQKVDGVDE
VAKKSKKEKDSKLEKALKQANDLIWNIKDELKKVCSNDLLEKILFNKQVGPVSGE
SAILRDVADGMSFGALLPCEEGSQGLVFSDDAYCYTGDTVAWTKVMVQTPNRKEM
TKTEFREISYLLKLLKKYKQDRIPEPETSASVAATPPSTASPAAYNSASADKPLSN
MKTITGLKLSRNDKEMKATIEKILGKLGTGTANKASLCISTKKEVKMNKKMEVDEKAN
IRVYSDGTSASQSKLQELFAHLSPWAEVKAEPVEVPAVRGKSGAALSKSKS
GOVKEGINKSEKRMKTLGGAADVDPGLSEHSAHVLEKGVVFSATLGLVDIVKGT
NSYKQLLEDENKRYMPTIDSGYEVAVKTLGNSNKLQEMPSKEDAIEHMKLFEKTN
AWHSNKFTYKPKFYPLEIDYGODEAVKTLNPGCTSKLKPQVODLTKMFDVSM
KAMVEYEDLQKMPGLKLSKRQIQAYSILSEVQAAVSGQSDSQIIDLNSRPTYL
PHDFGMKPLPLNADVAQSMLEMLNDLIEVAYSLLRGSDSDSDSKSDIPDVNKKL
TDIKVDVDRDSEAEIIRKYVKNTHATTHNAYDLEVIDFKIEREGCQPKYKFKQLHN
RULWHSRTNTPAGLSOGLRTAPPEAVPYGMFGKTYFADPMVSKSANSCTSQGD
PIGLILLEGVALNGMYELKHAHSISKLKPKGSHVSGKLGKTTFDPMVSKSANSCTSQGD
TGSSGSSNDTSLLYNEYIVYDIAQVNLKYLKLFNFKTSLW"
804.
old_sequence

```

old sequence

old_sequence

BASE COUNT	1048 a	847 c	1037 g
ORIGIN	73 bp upstream of AluI site		

Query Match	7.6%	Score 224.6	DB 9	Length 3795
Best Local Similarity	51.1%	Pred. No. 7.8e-47		
Matches 693	Conservative 11	Mismatches 612	Indels 39	Gaps 7
QY 1602	ctactatgtactccagatcattgaacaggatgaggctcgtgaatgctgaatgctgaatgcttaatttcgtaa	1661		
Db 1799	CTACTACAAGCTGCGAGCTTCCTGGAGGAGCGACAGAGAAACAGGATATTGGATATTCAGGTC	1858		
QY 1662	gtggggacgggttggaagtggaaaaattggaggccaaaaaactggaggagatgtccaaaaaac	1721		
Db 1859	CTGGGCGCGTGGGTACGGTGATCGGTAGCAACAACACTGGACACAGATCCGCTCCAGGA	1918		
QY 1722	tgaggcaatcaaggaaattcaaaagattattcttggagaagactggaaactcatcggaagc	1781		
Db 1919	GGATGCCATTGAGCACTTCATGAATTTATATGAAGAAAAAACCGGGAACGCTTGGCACTC	1978		
QY 1782	tggggaagttaaaaccaaatttcggaagcagcctggagagatttttaaccacttgaattgtga	1841		
Db 1799	C-----AAAAATTTCAGGAAGTATCCCAAAAGTTCTACCCCTCGGAGATTGA	2026		
QY 1842	ttatggtgttaagaagcaccac-----aaacggaaagatcatcagtgaaatgaaaaagttc	1895		
Db 2027	CTATGCCAGGATGAAGAGGCAGTGAAGAGCTGACAGTAATAATCTGGCACCACAGTCCAA	2086		
QY 1896	tcttgctoctcaattgctagaactcatgaagatgcttttcaatgtggagacatatagagc	1955		
Db 2087	GCTCCCAAGCCAGTTCAGGACCTCATCAAGATGATCTTTGATGTGAAAGTATGAAGAA	2146		
QY 1956	tgctatgtggaattgaaavtaatatgtcagaaaatgctctcttgggaactgaagcmaagrr	2015		
Db 2147	AGCCATGTGTGAGTATGAGATCGACCTTCAGAAGATGCCCTTGGGGGAAGCTGAGCAAAAG	2206		
QY 2016	aaatatggagaaaggttggaagcattaaactkkgtaacmrratttatttgaaggacaccg	2075		
Db 2207	GCAGATCCAGGCCGATACTCCATCTCAATGAGGTCCAGACGGCGGT-----GTCTCAG	2261		
QY 2076	ctnataagcactggctgtgttrgagaagcnaattgttgytsgagcmatsytttttc	2135		

RESULT	11
LOCUS	HUMADPPO
DEFINITION	Human placental poly(ADP-ribose) polymerase mRNA, partial cds.
ACCESSION	M17081
VERSION	M17081.1 GI:178151
KEYWORDS	.
SOURCE	Human placenta cDNA to mRNA, clone lambda-PAP803.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1320; 1560 to 1771) Suzuki.H., Uchida.K., Shima.H., Sato.T., Okamoto.T., Kimura.T. and

SOURCE Chinese hamster.
 ORGANISM Cricetus griseus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 Cricetus.
 REFERENCE 1 (bases 1 to 3047)
 AUTHORS Ganesh,A.N., Phillips,E., Shah,R., Affar,E.B., Poirer,G.,
 Thacker,J. and Meuth,M.
 TITLE Suppression of the radiation sensitive phenotype of hamster irs1
 and irs2 strains selected for resistance to 3 aminobenzamide
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3047)
 AUTHORS Ganesh,A.N. and Meuth,M.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-1999) Oncological Sciences, University of Utah,
 15 N. 2030 E., Salt Lake City, UT 84112, USA
 FEATURES
 source
 1. 3047
 /organism="Cricetus griseus"
 /db_xref="taxon:10029"
 /cell_line="v79-4"
 1. 3047
 /gene="PARP"
 1. 3042
 /gene="PARP"
 /EC_number="2.4.2.30"
 /codon_start=1
 /product="poly ADP-ribose polymerase"
 /protein_id="A045817.1"
 /db_xref="GI:5616520"
 /translation="MAEASERLYRYEYAKSGRASCKKSESIPKDSLRMAIMVQSPMF
 DKGPHWYHFDGSEVDFSELWDDOKVKTAEGAGVAGKGDGGS
 GRSKXTLGDFAEYAKSNRCKGOMEIKGOVPLSKMLDPKPOLGMDRWHP
 TCFVKREELGFPEYSASQLGFSLLSADREVLKQPLGVKSGKRGRDVRGADE
 VANKSKKDKDSKLEKALKQNDLWIKDELKAKCSTSLKELLIFNQOVPVSGE
 SAILRDVAGMAGFALLPKCESQLVFKSDAYCTGDTAWTKMVKQTQPSREKW
 TPREFREISYLKVKVKQDRIFPPETSAPEHPHPSPVTSAPTAVNSCPADKPLSN
 MKLTILGLKSQSKDEAKATEKLGKLTGSANNALCISTKEVERKMGKMEVOAAN
 VRYVDFDLQDAEAKSLQELLSHLSHLSWGAENVPEVAAPKSAAPSKSKG
 LYKEGVNKSXEMKLTLLGGAADVPDGLSEHSAHLEKGVFSATLGLDIVKGTN
 SYKQLLEDDEKSRWIFRSWGRVGTVIGSNKLEOMPSKEDAEVHFMLYEEKTGN
 WHSKNTFKPEFPLEIDYGDDEAVKLTVPKPTKSLPRAVELGVMIFDVPSMK
 KALVEYEDLQMLPKLSKQIQAAYSILSEVQVQSGSDSLDLNSRFTYLIIP
 HDEMKPDIPLNNADSVQAKVEMLDLIEVAYSLLRGSDSDSKDPTDVMYKLLKT
 DIKVDSDSEAEVIRKYVKNTHATPEYVLEMDIFKIEREGESORYKPKQLHNR
 RLHWGSRFTNFAGILSOGRLTAPPEAVTYGMEFGIYFADWVSKSNYCHTSQDP
 IGLILGEVALGNWYELKSHSLKPLKSHSVKGLGKTPDPSPASITILEGVEVPLGT
 GIPSGVNDTCLLYNEYIVYDIAQVNLKLLKLFNFKFTSLW"
 BASE COUNT 846 a 739 c 869 g 593 t
 ORIGIN
 Query Match 7.5%; Score 222.2; DB 10; Length 3047;
 Best local Similarity 50.4%; Pred. No. 3.2e-46;
 Matches 683; Conservative 14; Mismatches 619; Indels 39; Gaps 6;
 Qy 1602 ctactatgtactccatcattaaacaggatgaggtctgaagtctacgtatttcgttaa 1661
 Db 1701 CTATTACAGCTGCGAGCTTCGAGGATGACAAAGGAGCAGATCTCGATCTCCGGTC 1760
 Qy 1662 gtggggcgggtgtggagtgagaaaaattggaggggcaaaaactggaggagatgccaataac 1721
 Db 1761 CTGGGGCGAGTGGGCATGTCATTGGCAGTAAACAACTTGGACGATGCCATCCAAAGA 1820
 Qy 1722 tgaggcaatcaagaaatcaaaagattattcttgagaagactggaaactcattggaaagc 1781
 Db 1821 GGATGCTGTGAGCAGCTTCATGAACATATATGAAGAGAACACTGGCAATGCCCTGGCAGCT 1880
 Qy 1782 ttgggaatgaaacaaactttcggaagcagcctgggagattttaccacttgatgtga 1841
 Db 1881 -----AAAGAACTTCACAAAGTATCCCAAAAAGTTCTACCTCTGAGATTGA 1928
 Qy 1842 ttatgg-----tgttaagaagaacacaaacggaagatatcagtgaatgaaagtgc 1995

Db 1929 CTATGCCAGGATGAAGAGGCTGTAAGAAGCTGACGGGTAAACCTGGCACTAAGTCAAA 1988
 Qy 1896 tcttgcctctcaatgtctagaaactcaatgaagatgcttttcaatgttgagacatatagagc 1955
 Db 1989 GGTCCGGAAGGAGGAGTGCAGAGAGCTGTCGGGATGATCTTTGACGTGGAGAGCATGAAGA 2048
 Qy 1956 tctatgatggaattgaaawtaatatgtcagaataatgctctcttggaagcgaagcaggr 2015
 Db 2049 GGCCTTGGTGAATACGAGATTCACCTTCAGAAAGATGCCCTTGGGGAAGCTGAGCAAAAG 2108
 Qy 2016 aaatattgaggaagatttgaagcattaaactkrgtactmigrattttatttgaaggaacacg 2075
 Db 2109 GCAGATCCAGGCGCTACTCTACTCTTAGTGAGGTCCAG-----CAGGCAGTGTCT 2159
 Qy 2076 ctnatacaagcactgctgttgrgagaagcnaatgttgytgsagcmatsytttttc 2135
 Db 2160 CCAAGSCAGCAGTGCATCCAGATCTTCCAGATCTTTTCCAATCGATTCTACACCTTAATCCC 2219
 Qy 2136 actcttatcccttctattcatctcatattatcacgggatgaggtatttcatattcaaa 2195
 Db 2220 CATGACTTCGGAATGAAGAAGCCGCCACTCTCTGAACAAATGCAGACAGGGTGCAGGCCAA 2279
 Qy 2196 ggcgaataatgctgaagctctgcaggatatattgaaattgtcttcaa-----agatagtgg 2249
 Db 2280 GGTGGAGATGCTGGACAACCTCTCGACATCGAGGTGGCTATAGTCTTCTCAGGGGGGG 2339
 Qy 2250 ctctgataagcagctgataatctctctgtatgataataatataatgaactcactgtgaacat 2309
 Db 2340 GTCTGTATGATGACAGTAAGAGGCCCACTTGTATGATCAACTATGAGAAGCTCAAACTTGAT 2399
 Qy 2310 caccgcctgctcagcagtagtgaagattacaagttaattgagcagtagtctctcctcaaac 2369
 Db 2400 TTAGGTGTTGACCCGAGACTCTCGAAGAACTGAGGTCTATCAGGAAGTATTTTAAAGAACAC 2459
 Qy 2370 acatgctctactcacaaggactgctgcgtgaactggaggaggtttttcacttgatcg 2429
 Db 2460 TCATGCCACACACAAATGCTCTATGACCTGGAGGTGATGATATCTTTAAGATAGAGCG 2519
 Qy 2430 agatggagaacttaataagtaactcaagataataaataatctgcaataaagatgctatt 2489
 Db 2520 TGAGGAGAGA---GAGCCAGCGGTACAAACCTTCAAGCAGCTGCACAACTCGGAGGTTACT 2576
 Qy 2490 atggcacggttcaaggttgacaaatttggggaattcttagcaaggcgttaagaattgc 2549
 Db 2577 GTGGCAGCGGTCCAGGACCACTTTGCTGSCATCTCTGCACAGGCTCGCGGATAGC 2636
 Qy 2550 acctcctgaggcactgttactggttatgttcggaagcctctactttgcagatct 2609
 Db 2637 CCCACTGAAGCACCTGTGCACAGGCTACATGTTTGGAAAGGGATCTATTTGCTGACAT 2696
 Qy 2610 agtaagaagagcgcaataactgttattgtgtaggaataataactctgtagttgactgt 2669
 Db 2697 GGTGTCCAAAATGGCCAACTACTGTCATACATCTCAGGAGACCCGATGGCTGTGATACT 2756
 Qy 2670 tctttctgaggttcttaggagacatgtatgaactaaagaagccacgctccatgacaa 2729
 Db 2757 GTTGGGAGAAGTGCCTTGGAAACATGTACGAACATCAAGCATGGCTCACATATCAGCAA 2816
 Qy 2730 acctcaagaggggaagcattocaccaagggattagggcaaaacccgtgcacttggagtcaga 2789
 Db 2817 GTTACCAAGGGCAAGCACAGTGTCAAAGGCTTGGGCAAAAC---GAGCCCTGACCCCTC 2873
 Qy 2790 gttttagtgagtgaggatgctgttagttccctgcggcaagcgggtgcatcatcaat 2849
 Db 2874 GGCAGCATFACCTCTGGAGGGGTAGAGGTTCCACTGGGGACAGGGATCCCATCGGGTGT 2933
 Qy 2850 taggagctctgaactcaatgataatgataatgataatgataatgataatgataatgataat 2909
 Db 2934 TAATGACACCTCCCTTACATATATACGAGTACATCTCTATGACATTTGCTCAGGTGAACCT 2993
 Qy 2910 gcagttcttgctgaagggtgcgtttccatcaacaaga 2944

Db 2994 GAAATACCTGCTGAACACTCAAGTTAAATTTAAGA 3028

RESULT 13

RNPAP2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

RNPAP2 1939 bp DNA ROD 20-MAY-1992
R.norvegicus mRNA for poly(ADP-ribose) polymerase (clone pRATC).
X65497
X65497.1 GI:56849
NAD(+) ADP-ribosyltransferase.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

1 (bases 1 to 1939)
Thibodeau, J., Gradwohl, G., Dumas, C., Clairoux-Moreau, S., Brunet, G.,
Penning, C., Poirier, G.G. and Moreau, P.
Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase
catalytic domain and analysis of mRNA levels during the cell cycle
Biochem. Cell Biol. 67 (9), 653-660 (1989)
90027702
2 (bases 1 to 1939)
Potvin, F.
Direct Submission
Submitted (23-MAR-1992) F. Potvin, Molecular Endocrinology, CHUL
Research Center, 2705, Boul Laurier, Ste-Foy, Quebec, G1V 4G2,
CANADA

COMMENT

FEATURES

source

location/Qualifiers

1. .1939

/organism="Rattus norvegicus"

/strain="Sprague Dawley"

/db_xref="taxon:10116"

/chromosome="1q41-42"

/tissue_type="prostate"

/clone_lib="lambda EMBL4"

/clone="pRATC"

32. 1510

/EC_number="2.4.2.30"

/codon_start=1

/product="Poly(ADP-ribose) polymerase"

/protein_id="CAA46478.1"

/db_xref="GI:56850"

/db_xref="SWISS-PROT:P27008"

/translation="MKLTLKGAADPDPSLESHAVLPEKGGKVFSAATLGLVDIVKGT
NSYKIQLESKESRYWFRSNGRVGTIGSNKLEOMPSEDAVEHEMKLYEKTGN
AMHSKNTYKPKFHPPLAIDYGDDEAVKRLKVPKTKSLPKPQVQLVGMIFVESM
KKALVEYEDLQKMPGLKRRQIQAAYSILSEVQAVSQGSESOILDLNRFYILI
PHDFGKMKPLLDNTDSQVQKVEMLNDLIDIEVAYSLRGGSDSDSKDPIDVNYEKLK
TDIKVVDPSSEAEVIRKYVKNTHATNAYDLEVIDIFKIEREGESORYKPFRLHN
RLLWHSRTTFNAGILSGLRIAPPEAPVTGMFGKGIYFADMYKSANYCHTSGOD
PGLILLGEVALGNMYELKASHISKLPGKHSVGLGKTAPDPSASITLDGVEVPLG
TGIPSGVNDTCLLYNIYVDIAQYNLKLKLFNFTSLW"

CDS

524 a 468 c 521 g 426 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Score

219.8

DB 10

Length

1939

Indels

27

Gaps

5

1602 ctactgtactccagatcattgacaggatgatgggtctgagtgctacgtattctgtaa 1661
169 CTATTACAACTGCAACTTCTGGAGTCCGACAGAGAGAGAGGATTTGGATCTCCGGTC 228
1662 gtgggagcggttgggagtgagaaaattgaggggcaaaactggagagatgtcaaaaac 1721
229 CTGGGGCGGAGTGGGACAGTTATCGGCAGTTAACAAGCTTGAGCAGATGCGCTCCAAAGA 288
1722 tgaggcaatcaggaattcaaaagattattcttgaaagactgaaactggagagc 1781
289 GGATGCTGTTGAACACTTTCATGAAGTTATACGAAGAAGAACTGGAATTCCTGGCCTC 348

1782 ttgggaatgtaaaaccaaattttcgggaagcagcctgggagatttttaccacttgaattga 1841
349 GAAAAACTTCAAAAGTATCCAAAGAGTTCCACCC-----TCTGGCGATTGACTATGG 402
1842 ttatgggttaagaagac 1901
403 CCAGGACGAAGAGCAGCTAAAGAGCTGGGGTGAAGCCCTGGCACTAAGTGAAGCTGCC 462
1902 tctcaattgctagaactcatgaagatgcttttcaattgctggagacacacacacacacac 1961
463 GAAGCCGGTGGAGAGCTCGTGGGATGATCTTCGACGTGGAGAGCATGAAGAAGGCGCTT 522
1962 gatgaatttgaaawtaatatgtcagaaatgcctctctgggaagcgaagtaagmagraatat 2021
523 GGTGGAGTACGAGATTGACCTTTCAGAAAGATGCCCTTGGGAAACTGAGCAAGGCGAGAT 582
2022 tgagaaagattgaagcattacactkgrtacmoratttttgaagacacacacacacacac 2081
583 CAGGCCGCCCTACTCTATCTCAGTGGAGTCCAG-----CAGCAGTGTCCCAAGG 633
2082 aagcactggctgttrgagaagcnaaattgtgtgsgagcmatsytttttcaactctt 2141
634 CAGCAGTGAATCTCAGATCCTGGATCTCTCAATCGCTTCTACACCCCTCATCCCCCATGA 693
2142 atcccttcttctcctcctcattatacagggatgaggtatttcaattcattcaaaagcgaa 2201
694 CTTTGAATGAAGAAGCCGCCACTTCTAGACAACACACAGACAGTGTGAGGCCAAGGTGGA 753
2202 aatgcttgaagcctctcagagatttgaaattgcttcaagata-----gttggtctga 2255
754 GATGTGGACAACACTCTCTGGAGATCGAGGTTCCTTATAGTCTTCTCAGGGGTGGTCTGA 813
2256 tagcgacagtgtgaatctcttgatgataaataataataataataataataataataata 2315
814 TGACAGCAGTAGAGCAGCCCATCGATGTCACTATGAGAAGCTCAAACTGACATTAAAGGT 873
2316 gctggctcacgatatgaaagattcaaaagtaataatgagcagatctcctcaacacacacac 2375
874 GGTGACAGAGATTCTGAAGAGGCTGAGGTCTATCAGGAAGTATGTAAGAAGACACTCAGC 933
2376 tctactcaacagagactggtcgtggaagctgttttcaacttgatcgagatgg 2435
934 GACCAGCCAAATGCTATGACCTCGAGTGTATGACATCTTTAGATAGAGGAGAGGG 993
2436 agaaacttaaatgactcaagataataaataataatcgcatacaacaagatgctattatgca 2495
994 AGA---GAGCCCAAGCGTACAAGCCCTTCAGGCAGCTTCAACACCCGAGAGACTGCTGTGGA 1050
2496 cgggtcaaggttgacgaattttgtgggaattcttagtcaaggctaaagaattgacactcc 2555
1051 CGGGTCCAGGACCAACAACTCGAGGCATCTGTCAAGGGTCTGCGGATAGCCCAACC 1110
2556 tgaggcacctgttactggctatatgttcggcaagcctctacttgcagatctagtagaag 2615
1111 TGAAGCAGCTGTGACAGGCTACATGTTTGGGAAAGAAATCTACTTGTGTGATGTTGTC 1170
2616 caagagcgcaataactgttattgtggatgaggaataatcctgtaggtttgagttctttc 2675
1171 CAAAGTGGGAATCTATGTCACACAGCTCTCAGGAGAGACCGGATTTGCTTAATATGTTGG 1230
2676 tgaggttcttagagacatattgaataaagaacacacacacacacacacacacacacac 2735
1231 AGAAGTTGGCTTGGAAACATGTACCAACTCAAGCATGTTCTCACATCAGCAAGTTACC 1290
2736 aagagggaagcattcgac 2795
1291 AAAGGCAAGCAGACAGTGTCAAGAGTTTGGGCAAAACCG---CCCTGACCTTCGCCAG 1347
2796 gaagtggaggtatgagtctgtagttccctgcggcaagcgggtgccaatcaattaggaag 2855
1348 CATCACCTTGGATGGTGTAGAGGTTCCTGGGAACAGGATTCCTGCTGGTGTAAATGA 1407

QY 2856 ctctgaactcatgatacagtagtactatgctctacacacacatccagggtgaagatgcagt 2915
 Db 1408 CACCTGCGCTGCTGTATACGAGTACATGCTTACGACATTCCTCAGGTGATCTGAAGTA 1467
 QY 2916 ctgtctgaaggtgcgtttccatcacacaaga 2944
 Db 1468 CCTACTGAAACTGAAGTTCATTAATTTCAAGA 1496

 RESULT 14
 LOCUS AX058340 3045 bp DNA PAT 17-JAN-2001
 DEFINITION Sequence 24 from Patent WO0071719.
 ACCESSION AX058340
 VERSION AX058340.1 GI:12310800
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3045)
 Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.
 Human poly(adp-ribose) polymerase 2 materials and methods
 Patent: WO 0071719-A 24 21-DEC-2000;
 ICOS CORPORATION (US)
 FEATURES
 Location/Qualifiers
 1..3045
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..3045
 /codon_start=1
 /protein_id="CAC22488.1"
 /db_xref="GI:12310801"
 /translation="MAESDRLRYVEYAKSGRACKKCSSEIPKDSLRNAINVQSPMF
 DGKVPWHYHFCSPKVGHSIRHPDVFSELRWDDQKVRKKTAEAGVVGKQDGI
 GSKAEKTLGDFRAEYAKSNRSTCKGCMERIEKQVRLSKKMDPEKPOLGMDRWYHP
 GCVNKEELGDFPEYSOLQFSLATEDKALKKOLPGVSKSEKRGKDEVDGVE
 VAKKSKKEDKSKLEKALKQNDLWIKNDLKKVCSNDLKKELLIPKQOVPSE
 SAILEDVADGVGFLPCBECQSGOLVFVSDAYYCTGVTAWTKMKWKTOTPNREK
 TPEFREISVLLKLVKVKDRIFPPETSASVAATPPSPASAPAAVNSASADKPLSN
 MKTLTGKLSRNKDEVKAMTEKLGKLTGTANKASLCISTKKEVEKMKMEVEKEAN
 IRVSEDFLODVSAKSLQELFLAHILSPWGAKEVPAEVPVPRGSGALSKSKK
 QVKEGINKSRMKTLLKGGAAVDPPDGLSEHSAHLEKGGKVFSAITGLVDIVKGT
 NSYIKQLLEDDEKKNRYIFRSWGRVTVIGSNKLEQMPKEDAEQKMKLYEEKTN
 AHSNKETRYPKKFFLEIDYQDDEAVKLVNPTGSKLPKQVQDLKMTFVDSM
 KRAMVEYEDLOKPLGKLSKQIOAAVSLSEVOQAVSOGSSDLSNRFYTLI
 PHDFGKKPPLNNADSVQKVMELNDLIDIEVAYSLRGGSDSSKDPIDNVEKLL
 TDIKVDRSEAEIIRKYKNTHTTHSAYDLEVIDIFKIEREGECORYKPKQLHN
 RLLWHGSRITNFAIGLSQGLRIAPPEAPVTVMFGKGIYFADMYSKSANYHTSQGD
 PIGLILLGEVALGNMYELKHASHISRLPKGHSVRKGLGKTPDPSPANISLDGVDVPLG
 TGSSGVIDTSLLYNEYIVDYIAOVNLYLLKLFKFNKLSLW"
 BASE COUNT 863 a 710 c 851 g 621 t
 ORIGIN

 Query Match 7.5%; Score 219.8; DB 6; Length 3045;
 Best Local Similarity 50.9%; Pred. No. 1.3e-45;
 Matches 690; Conservative 11; Mismatches 615; Indels 39; Gaps 7;

 QY 1602 ctactgtactccagatcatgaacaggatgatgggtctgagtgctagctatcttcgtaa 1661
 Db 1704 CTACTAAGCTTCAGCTCTCTGGAGCAGCAGCAAGGAAAAACAGGTATTTGGATTATTCAGGTC 1763.
 QY 1662 gtggggacgggttgaggatgagaaattggaggcgcaaaactggaggagatctcaaaaac 1721
 Db 1764 CTGGGCGGTGCTGGGTACCGGTGATCGGTAGCAACAACATGGACAGATGCCCTCCAGGA 1823
 QY 1722 tgaggcaatacgaagattcctaagaattattttttgagaagactggaaactcatgggaagc 1781
 Db 1824 GGATGCCATTGAGCAGCTTCTGAAATATATGAAGAAAAAACCAGGAGCGCTTGGCAGCTC 1883
 QY 1782 ttgggaatgaaacaaatttttcggagcagcagctgggagattttaccacactgtatgta 1841

Db 1884 C-----AAAAATTCAGAAAGTATCCAAAAAGTTTTTACCCCTCGAGATTGA 1931
 QY 1842 ttatggtgttaagaaagcacca-----aaacggaagatacagtgaaatgaaaagttc 1895
 Db 1932 CTATGGCCAGGATGAAGAGGCGAGTGAAGAAGCTCAGAGTAAATCTCTGGCACCAGATCCAA 1991
 QY 1896 tctgtctctcaattctgagaactcatgaagatgcttttcaattgagtgagacatatagagc 1955
 Db 1992 GCTCCCAAGCCAGTTCAGAGACCTCATCAAGATGATCTTTGATGTGGAAGATGATGAAGA 2051
 QY 1956 tgcctatgagaaatttgaaataatgtaacaaatgcctcttggaagcctaagcmaggr 2015
 Db 2052 AGCCATGCTGGAGTATGAGATCGACCTTCAGAAAGATGCCCTTGGGGAAGCTGACGAAGA 2111
 QY 2016 aaatatgagraggatttgaagcattaaactkrgtctacmrratttatttgaaggacacog 2075
 Db 2112 GCAGATCCAGGCGCATACTCCATCTCAGTGAGGTCCAGCAGGCGGT-----GTCTCAG 2166
 QY 2076 ctnataagcactggctttrgagaagcctnaattgttgytgsagcmatsyttttc 2135
 Db 2167 GGCAGCAGGACTCTCAGATCCTGGATCTCTCAAAATCGCTTTTACACCTGTATCCCCAC 2226
 QY 2136 actcttactccttctattctcattatatacagggatgaggtgatttcattatcaaa 2195
 Db 2227 GACTTTGGGATGAAGAAGCCTCCGCTCCTGAACAATGCAGA---CAGTGTCCAGGCCAA 2282
 QY 2196 ggcgaatactgtaagctctcagagatatgaaattgttctca-----agatagttgg 2249
 Db 2283 GGTGAAATGCTTGACAACTCTGCACATCGATGTCACTATGAGAAGCTCAAAACTGACAT 2402
 QY 2250 ctctgatagcagcagtgatgaactctctgataataataataatactgataaactcactgacat 2309
 Db 2343 GTCTGATGATAGCAGCAAGGATCCCATCGATGTCACTATGAGAAGCTCAAAACTGACAT 2402
 QY 2310 caccocgtggtcagcagtagtgaagattacaagtttaattgagcagtagtctctcctcaac 2369
 Db 2403 TAAGGTGGTGTGACAGAGATTCTGAAGAAGCCGAGATCATCAGGAAGTATGTTAAGAACA 2462
 QY 2370 acatgctctactcaacaggactgctgctggaactgaggaagtgttttcaactgacg 2429
 Db 2463 TCATGCAACACACACAGTCGGCTATGACTTGGAGTCACTGATATCTTTAAGATAGACG 2522
 QY 2430 agatggaaacttaataagactactcaagataaaaaataatctgcatacaacagatgctatt 2489
 Db 2523 TGAAGCGGATGCCAGCTTACACCCCTTTAAGCAG---CTTCATAACCGAAGATTGCT 2579
 QY 2490 atggcacggttcaaggttgaagaattttgtgggaattcttagtaaggcgttaagaattgc 2549
 Db 2580 GTGGCAGCGGTCCAGGACCACCAACTTTGCTGGGATCTGTCCCAGGCTCTTCGGATAGC 2639
 QY 2550 acctctgagcagcctgttactggtctatatgttcggcaaaagcctctacttgcagatct 2609
 Db 2640 CCGGCTGAAGCGCGCGGTGACAGGCTACATGTTGGTAAAGGGATCTATTTCGCTGACAT 2699
 QY 2610 agtaagcagagcgcaactactgttatgtgtagagaataatctctgtagttttagct 2669
 Db 2700 GGTCTCCAAGAGTGCACACTACTACCATCTCTCAGGGAGACCCCAATAGCTTTAATCCT 2759
 QY 2670 tcttctgaggttctttaggagacatgtatgaactaaagaagccacgtccatggaca 2729
 Db 2760 GTTGGGAGAAAGTTGCCCTTGGAAACATGTATGAACGAAGCAGCTTCACATATCAGCAG 2819
 QY 2730 acctccaagagggaagcattcgaccaagggattagggcaaaacccgtgcactggagtcaga 2789
 Db 2820 GTTACCCAGGCGAAGCAGAGTGTCAAGAGTTTGGGCAAAACTACCCCTGATCCTTCAG- 2878
 QY 2790 gtttgtgaagtggaggatgatgctagttccctgcgcaagccggtgcacatcaat 2849
 Db 2879 --CTAACATTAGTCTGGATGCTGTAGAGCTTCTCTTTGGGACCGGGATTTCACTGGTGT 2936
 QY 2850 tagagctctgaactcatgatacagtagtactatgatacaacacatcccagggtgaagat 2909

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 14:35:10 ; Search time 298.35 Seconds
(without alignments)
8474.115 Million cell updates/sec

Title: US-09-236-995D-1
Perfect score: 2949
Sequence: 1 atgcggcgccgcaagc.....gtttccatcacagagtag 2949

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2935.8	99.6	2949	20	Maize poly ADP-rib
2	2796	94.8	3212	21	DNA encoding the p
3	2720	92.2	3211	21	DNA encoding the p
4	224.6	7.6	3793	17	DNA-binding domain
5	221.8	7.5	3580	20	Human poly(ADP-rib
6	219.8	7.5	3045	22	Human tankyrase2 r
7	205.2	7.0	3200	22	Fusion protein PAR
8	204.4	6.9	2147	21	DNA encoding the p
9	200.2	6.8	2295	21	DNA encoding the p
10	191.8	6.5	1566	22	Human poly(ADP-rib
11	191.8	6.5	1814	22	hparp2 cDNA. Homo

12	191.8	6.5	1843	21	AAZ44287	Human brain PARP2
13	191.8	6.5	1843	22	AAC82090	Human brain poly-A
14	191.8	6.5	1912	21	AAC77118	Human ORFX ORF2673
15	185.4	6.3	1980	22	AAH14941	Human cDNA sequenc
16	171	5.8	864	22	AAC85320	3' end fragment of
17	160	5.4	5345	16	AAT04221	5.3 kb EcoRI fragm
18	158.4	5.4	2682	16	AAT04219	Poly(ADP-ribose)-p
19	131	4.4	1592	16	AAT04220	Poly(ADP-ribose)-p
20	127.6	4.3	936	22	AAF58252	Oligonucleotide D1
21	127.6	4.3	936	22	AAF58254	Oligonucleotide D1
22	127.6	4.3	936	22	AAF58257	Oligonucleotide D1
23	127.6	4.3	936	22	AAF58259	Oligonucleotide D2
24	127.6	4.3	936	22	AAF58262	Oligonucleotide D2
25	127.6	4.3	938	22	AAF58255	Oligonucleotide D1
26	125	4.2	595	21	AAZ80132	Human colon cancer
27	119.4	4.0	936	22	AAF58252	Oligonucleotide D1
28	119.4	4.0	936	22	AAF58254	Oligonucleotide D1
29	119.4	4.0	936	22	AAF58257	Oligonucleotide D1
30	119.4	4.0	936	22	AAF58259	Oligonucleotide D2
31	119.4	4.0	936	22	AAF58262	Oligonucleotide D2
32	119.4	4.0	938	22	AAF58255	Oligonucleotide D1
33	68.6	2.3	396	22	AAH42823	Nucleotide sequenc
34	63.6	2.2	3308	22	AAF63990	Human tankyrase2 e
35	63.4	2.1	395	22	AAF66086	Novel human polynu
36	55.6	1.9	4590	22	AAH24065	Yeast AOD9604-asso
37	50.4	1.7	416	22	AAC85306	EST R28562, fragme
38	46	1.6	244	22	AAF58238	Oligonucleotide D1
39	45.8	1.6	1143	22	AAC85314	5' end fragment of
40	43.4	1.5	244	22	AAF58238	Oligonucleotide D1
41	41	1.4	1921	19	AAV16949	Nucleic acid encod
42	40.8	1.4	550	22	AAC85318	Fragment P2-1 derl
43	40.8	1.4	550	22	AAC85319	Fragment P2-9 derl
44	40.4	1.4	14041	22	AAH48024	Internal control B
45	39.4	1.3	1473	20	AAH80867	Degenerate DNA seq

ALIGNMENTS

RESULT 1

AAH89542
ID AAX89542 standard; cDNA: 2949 BP.

XX AC AAX89542;

XX DT 06-OCT-1999 (first entry)

XX DE Maize poly ADP-ribose polymerase gene.

XX KW PARP; poly ADP-ribose polymerase; metabolism; maize; transform; ss.

XX OS Zea mays.

XX FH Key Location/Qualifiers

FT CDS 1..2949

FT FT /*tag= a

FT FT /product= "PARP"

FT FT /transl_except= (pos:1453..1455, aa:Xaa)

FT FT /transl_except= (pos:1495..1497, aa:Xaa)

FT FT /transl_except= (pos:1504..1506, aa:Xaa)

FT FT /transl_except= (pos:1534..1536, aa:Xaa)

FT FT /transl_except= (pos:1537..1539, aa:Xaa)

FT FT /transl_except= (pos:1561..1563, aa:Xaa)

FT FT /transl_except= (pos:1582..1584, aa:Xaa)

FT FT /transl_except= (pos:1588..1590, aa:Xaa)

FT FT /transl_except= (pos:1594..1596, aa:Xaa)

FT FT /transl_except= (pos:1975..1977, aa:Xaa)

FT FT /transl_except= (pos:2011..2013, aa:Xaa)

FT FT /transl_except= (pos:2014..2016, aa:Xaa)

FT FT /transl_except= (pos:2026..2028, aa:Xaa)

FT FT /transl_except= (pos:2047..2049, aa:Xaa)

FT FT /transl_except= (pos:2050..2052, aa:Xaa)

/transl_except= (pos:2053..2055, aa:Xaa)
 /transl_except= (pos:2056..2058, aa:Xaa)
 /transl_except= (pos:2077..2079, aa:Xaa)
 /transl_except= (pos:2095..2097, aa:Xaa)
 /transl_except= (pos:2107..2109, aa:Xaa)
 /transl_except= (pos:2116..2118, aa:Xaa)
 /transl_except= (pos:2119..2121, aa:Xaa)
 /transl_except= (pos:2122..2124, aa:Xaa)
 /transl_except= (pos:2125..2127, aa:Xaa)
 /transl_except= (pos:2128..2130, aa:Xaa)
 /note= "Xaa = unknown"

W09937789-A1.

29-JUL-1999.

26-JAN-1999; 99WO-US01591.

27-JAN-1998; 98US-0072785.

(PION-) PIONEER HI-BRED INT INC.

Mahajan P, Zuo Z;

WPI; 1999-444613/37.

P-PSDB; AAY28464.

New maize poly ADP-ribose polymerase gene useful in transforming plants to alter their metabolic state

Claim 1; Page 30-35; 48pp; English.

The sequence is the maize poly ADP-ribose polymerase (PARP) gene. The sequence codes a PARP with 982 amino acids (AAY28464). PARP is required in the cell in most cases of DNA repair, recombination, rearrangement and transposition. PARP gene and antisense gene can be used to transform plant cells and alter the metabolic state of the transformed cell. This is useful in enhancing disease resistance in plants and methods of genetic transformation of plants. Plants transformed with either a sense or antisense PARP nucleotide sequence may be utilized to increase transformation frequency in plant cells. The enzyme also plays a role in cellular stress, so may be beneficial for prevention of plant disease or pathogen attack.

Sequence 2949 BP; 895 A; 539 C; 732 G; 756 T; 27 other;

Query Match 99.6%; Score 2935.8; DB 20; Length 2949;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2949; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcgccgcgcgaaggcggtggaagcgagatgccaagtctggcgccgctctgtgc 60
 |||||
 DB 1 atggcgccgcgcgaaggcggtggaagcgagatgccaagtctggcgccgctctgtgc 60
 QY 61 aagtcagtcggtccctctatgcgaaggaccagctccgtctgtggcaagatggttcaggcg 120
 |||||
 DB 61 aagtcagtcggtccctctatgcgaaggaccagctccgtctgtggcaagatggttcaggcg 120
 QY 121 tcacagtcagcgtcttcagtcgagatgtgaacctgaccaggtgcattcttcagcaagaag 180
 |||||
 DB 121 tcacagtcagcgtcttcagtcgagatgtgaacctgaccaggtgcattcttcagcaagaag 180
 QY 181 aaccagataaaatccggtgacgatgttgaaggatagatgcacttagatgggatgatacaa 240
 |||||
 DB 181 aaccagataaaatccggtgacgatgttgaaggatagatgcacttagatgggatgatacaa 240
 QY 241 gagaagatacgaactacgtgtggagtgctcagctggtgataagttctcagctgctct 300
 |||||
 DB 241 gagaagatacgaactacgtgtggagtgctcagctggtgataagttctcagctgctct 300
 QY 301 ccgagaaaatgtacaattgagatgtccatctgcccgtacttcatgtagacgagtcagtc 360
 |||||

DB 301 cctgagaaaatgtacaattgagattgctccatctgcccgtacttcatctagatagatgcagtc 360
 QY 361 gaaaagattacaaaagatcggtccggtcttccagcttaagcttgagagtgaaaggtcccaag 420
 |||||
 DB 361 gaaaagattacaaaagatcggtccggtcttccagcttaagcttgagagtgaaaggtcccaag 420
 QY 421 ggtataccatggtatcatgccaaactgtttctttgaggtatccccctctcgtcaactgttgag 480
 |||||
 DB 421 ggtataccatggtatcatgccaaactgtttctttgaggtatccccctctcgtcaactgttgag 480
 QY 481 aagttctcaggtgggatactttgtccgagtaggataagagaaacccatgctcgatctgtt 540
 |||||
 DB 481 aagttctcaggtgggatactttgtccgagtaggataagagaaacccatgctcgatctgtt 540
 QY 541 aaaaagatgttggcaacaataaacaataaggttccaaagcgaagaaagtgaaaaat 600
 |||||
 DB 541 aaaaagatgttggcaacaataaacaataaggttccaaagcgaagaaagtgaaaaat 600
 QY 601 gatattgatgctacaaaatcccaggtttagatgaaagtacatctcgaaggtacagtcgga 660
 |||||
 DB 601 gatattgatgctacaaaatcccaggtttagatgaaagtacatctcgaaggtacagtcgga 660
 QY 661 acaaaaggcaactgttagaccccagtcggttccaaatactagttcagctgatatccaaacta 720
 |||||
 DB 661 acaaaaggcaactgttagaccccagtcggttccaaatactagttcagctgatatccaaacta 720
 QY 721 aagcttaaggagcaaaagtacacacactttggaagttaaaggatgaacttaagactaatga 780
 |||||
 DB 721 aagcttaaggagcaaaagtacacacactttggaagttaaaggatgaacttaagactaatga 780
 QY 781 tcggctgctgaattaaaggatagcttggagcctaagtggcaggatcacatcaggaccagaa 840
 |||||
 DB 781 tcggctgctgaattaaaggatagcttggagcctaagtggcaggatcacatcaggaccagaa 840
 QY 841 aggcacctattgagtcgctgtgcgagatggaatgctatttgagcgtggttctctgcccga 900
 |||||
 DB 841 aggcacctattgagtcgctgtgcgagatggaatgctatttgagcgtggttctctgcccga 900
 QY 901 gctctgctaaatggcatgactattataatggtcagtcacacatgcagtcggttaagtgtca 960
 |||||
 DB 901 gctctgctaaatggcatgactattataatggtcagtcacacatgcagtcggttaagtgtca 960
 QY 961 gagtgtccaaagtgtacatactctgtccacagaacctgtccgctgtaagaagaagtggcaa 1020
 |||||
 DB 961 gagtgtccaaagtgtacatactctgtccacagaacctgtccgctgtaagaagaagtggcaa 1020
 QY 1021 attccacatggaacaaagaatgattacattgaagtgttcaaatctcctcaaaagttaag 1080
 |||||
 DB 1021 attccacatggaacaaagaatgattacattgaagtgttcaaatctcctcaaaagttaag 1080
 QY 1081 aaaccagagaggttcttccccaatgtccactgagaaatctggaagttaaagcaactcag 1140
 |||||
 DB 1081 aaaccagagaggttcttccccaatgtccactgagaaatctggaagttaaagcaactcag 1140
 QY 1141 agaacatcattgctgtcttcttaaggggttgataaaaattaaaggtttctgttggtagcaaa 1200
 |||||
 DB 1141 agaacatcattgctgtcttcttaaggggttgataaaaattaaaggtttctgttggtagcaaa 1200
 QY 1201 tcaaaagaagcagcaaaaatgagtggaatgagagctcaaaactgtggtgccaactctat 1260
 |||||
 DB 1201 tcaaaagaagcagcaaaaatgagtggaatgagagctcaaaactgtggtgccaactctat 1260
 QY 1261 gccaggtgtgtcacaagatattgattgtttaaattgcattggtgagctcgacaatgaaaaat 1320
 |||||
 DB 1261 gccaggtgtgtcacaagatattgattgtttaaattgcattggtgagctcgacaatgaaaaat 1320
 QY 1321 gctgaagtcaagaaagcaagggaggtgaagataccaattgtaaggggaggttcatctgga 1380
 |||||
 DB 1321 gctgaagtcaagaaagcaagggaggtgaagataccaattgtaaggggaggttcatctgga 1380
 QY 1381 gaattgttcaaaagcaaaaatgctccattgattgattataaactggaatgccttagag 1440
 |||||
 DB 1381 gaattgttcaaaagcaaaaatgctccattgattgattataaactggaatgccttagag 1440

Qy	721	aagcttaagagcaaaagtgaacacacttggaaagttaaaggatgacttaagactcatgta	780
Db	801	aagcttaagagcaaaagtgaacacacttggaaagttaaaggatgacttaagactcatgta	860
Qy	781	tcggctctgaattaaaggatgctctgagcgctaagtggcaggagatacatcaggaccagaa	840
Db	861	tcggctctgaattaaaggatgctctgagcgctaagtggcaggagatacatcaggaccagaa	920
Qy	841	aggcacctattggtatcgctgtcggtatgaaatgctatttggagcgctgggtccttgccca	900
Db	921	aggcacctattggtatcgctgtcggtatgaaatgctatttggagcgctgggtccttgccca	980
Qy	901	gtctgtgctaaatggcagtactattataatggtcagtaaccaatgcagtggttaatgtctca	960
Db	981	gtctgtgctaaatggcagtactattataatggtcagtaaccaatgcagtggttaatgtctca	1040
Qy	961	gagtggtccaaagtgtacatactctgccacagaacctgtcccggttaagaagaagtggcaa	1020
Db	1041	gagtggtccaaagtgtacatactctgccacagaacctgtcccggttaagaagaagtggcaa	1100
Qy	1021	attccacatggaaacaaagaattgattacettatgaaagtgtgttccaaatctcaaaaggttaag	1080
Db	1101	attccacatggaaacaaagaattgattacettatgaaagtgtgttccaaatctcaaaaggttaag	1160
Qy	1081	aaaccagagaggtctctccaccaatgtccactgagaatctggaagttaagaagaactcag	1140
Db	1161	aaaccagagaggtctctccaccaatgtccactgagaatctggaagttaagaagaactcag	1220
Qy	1141	agaacatcatgtctctctaaagggttgataaaattaaagttttctgtgttaggacaa	1200
Db	1221	agaacatcatgtctctctaaagggttgataaaattaaagttttctgtgttaggacaa	1280
Qy	1201	tcaaaagaagcagcaaaatgagtggaattgagaagctcaaaacttgctggtgccacttctat	1260
Db	1281	tcaaaagaagcagcaaaatgagtggaattgagaagctcaaaacttgctggtgccacttctat	1340
Qy	1261	gccaggggtgtcaaaagattgattgttttaattgcatgtggtgagctcgacaatgaaatt	1320
Db	1341	gccaggggtgtcaaaagattgattgttttaattgcatgtggtgagctcgacaatgaaatt	1400
Qy	1321	gctgaagtcaggaagaagaaggagctggaagataccaattgtaaagggaggggttaccattgga	1380
Db	1401	gctgaagtcaggaagaagaaggagctggaagataccaattgtaaagggaggggttaccattgga	1460
Qy	1381	gaatgtgtt-aaagaacaaaatgctgcatttgaattgttataaac- -ggaatgacctta	1437
Db	1461	gaatgtgttataaaagaacaaaatgctgcatttgaattgttataaacatagagaatgacctta	1520
Qy	1438	gagctctcaaaaggcmgtactgtcactgttaaaagttaaggccgaagtgtctgttctcatya	1497
Db	1521	gagctctcaaaaggcagtagtactgtcactgttaaaagttaaggccgaagtgtctgttctcaty	1578
Qy	1498	agtcctcctggtttgaagaataactcacaactctctatratgtggaaaagcatacaca	1557
Db	1579	agtcctcctggtttgacaag-atactgtctacatt-cttgaagatgggaaaagcatacaca	1636
Qy	1558	tgcacmctttaaactgttctgacctgncacnaggtgtgyacaggtcactactgtactccag	1617
Db	1637	tgcaaccttaaacatg-tctgacctggcactagggtgtgaaaca-gctactatgtactccag	1694
Qy	1618	atcaattgaacagatgattgggtctgagtgctacgtatttctgaagtgggagcaggggttggg	1677
Db	1695	atcaattgaacagatgattgggtctgagtgctacgtatttctgaagtgggagcaggggttggg	1754
Qy	1678	agtgagaaaaattggagggcacaaaactggaggagatgtcaaaaaactgagccaatcaaggaa	1737
Db	1755	agtgagaaaaattggagggcacaaaactggaggagatgtcaaaaaactgagccaatcaaggaa	1814
Qy	1738	ttcaaaagattattcttttgaaagactggaactcatgggaagcttggggaattgaaaacc	1797
Db	1815	ttcaaaagattattcttttgaaagactggaactcatgggaagcttggggaattgaaaacc	1874


```
QY 1798 aatttcggaagcagcctggagagattttaccacacttgatgttgaattatggtgttaagaaa 1857
Db 1875 aatttcggaagcagcctggagagattttaccacacttgatgttgaattatggtgttaagaaa 1934
QY 1858 gcacaaaacggaagatatacagtgaaatgaaagtctcttctcctcaattgctagaa 1917
Db 1935 gcacaaaacggaagatatacagtgaaatgaaagtctcttctcctcaattgctagaa 1994
QY 1918 ctcatgaagatgcttttcaatgtggagacatatagagctgctatgatggaatttgaawt 1977
Db 1995 ctcatgaagatgcttttcaatgtggagacatatagagctgctatgatggaatttgaawt 2054
QY 1978 aatattgcagaatgctcttctggaagcctaagcmagrraaataattgaaagatttga 2037
Db 2055 aatattgcagaatgctcttctggaagcctaagcmagrraaataattgaaagatttga 2114
QY 2038 gcattaactkrgrtcmgratttatttgaaggaacccgctnatacagacactgctgttir 2097
Db 2115 gcattaactgagatacagaattta-ttgaaggaacccgctgatcaagcactggc-tgtta 2172
QY 2098 gagaagctnaattgttgytsgagcmatssytttttcaactcttattccttctattc 2157
Db 2173 gagaagctnaattgttgytsgagcactgcttttcaactcttattccttctattc 2232
QY 2158 ctcatattatcggatgagatgatttcatatttcaataaagcggaataatgcttgaagctcg 2217
Db 2233 ctcatattatcggatgagatgatttcatatttcaataaagcggaataatgcttgaagctcg 2291
QY 2218 caggatattgaaatgcttcaagatagttgcttcgtagcgacagtgatgaatctct 2277
Db 2292 caggatattgaaatgcttcaagatagttgcttcgtagcgacagtgatgaatctct 2351
QY 2278 gatgataatataataaacttcaacttgacatacacccccctggctcacgatagtgaagat 2337
Db 2352 gatgataatataataaacttcaacttgacatacacccccctggctcacgatagtgaagat 2411
QY 2338 tacaagttaattgagcagtatctctcaacacacacatgctcttactcacaagagactgctg 2397
Db 2412 tacaagttaattgagcagtatctctcaacacacacatgctcttactcacaagagactgctg 2471
QY 2398 ctggaaactggaagaaagtttttcaacttgatcgagatggagaaacttaataagtaactcaaga 2457
Db 2472 ctggaactggaagaaagtttttcaacttgatcgagatggagaaacttaataagtaactcaaga 2531
QY 2458 tataaaaataatctgcatacaagaatgctattatgagcaggttcaaggttgacgaatttt 2517
Db 2532 tataaaaataatctgcatacaagaatgctattatgagcaggttcaaggttgacgaatttt 2591
QY 2518 gtgggaattcttagtcaagggctaaagaattgcaactcctctgaggaacctgttactggctat 2577
Db 2592 gtgggaattcttagtcaagggctaaagaattgcaactcctctgaggaacctgttactggctat 2651
QY 2578 atgttcgcaaaagcctctactttgcagatctagtaagcaagcgcaataactgttat 2637
Db 2652 atgttcgcaaaagcctctactttgcagatctagtaagcaagcgcaataactgttat 2711
QY 2638 gtggataggaataaactctgtagttgtgcttcttcttctgaggttgccttaggagacatg 2697
Db 2712 gtggataggaataaactctgtagttgtgcttcttcttctgaggttgccttaggagacatg 2771
QY 2698 tatgaactaaagaaacgacgtccatggaacaaactcccaaggggaagcattcgaccaag 2757
Db 2772 tatgaactaaagaaacgacgtccatggaacaaactcccaaggggaagcattcgaccaag 2831
QY 2758 ggaattaggcaaaacccgtgcactggagtcagagatttgaagtggagggatgctgcga 2817
Db 2832 ggaattaggcaaaacccgtgcactggagtcagagatttgaagtggagggatgctgcga 2891
QY 2818 gttccctgcggaacggttgcatacatcaattagagctctgaactcatgtacaatgag 2877
Db 2892 gttccctgcggaacggttgcatacatcaattagagctctgaactcatgtacaatgag 2951
QY 2878 tacatgctgtacaaacacatcccaagtggaagatgcaggttcttctgctgaaggtgcttccat 2937
```

```
Db 2952 tacatgctgtacaaacacatcccaagtggaagatgcaggttcttctgctgaaggtgcttccat 3011
QY 2938 cacaagaggtag 2949
Db 3012 cacaagaggtag 3023
```

```
RESULT 3
AAZ60615
ID AAZ60615 standard; DNA; 3211 BP.
XX
AC AAZ60615;
XX
DT 16-MAY-2000 (first entry)
XX
DE DNA encoding the poly(ADP-ribose) polymerase ZAP1 protein of Zea mays.
XX
KW ZAP1; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
KW pest; drought; heat; fungi; nematode; seed-shatter; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS l13..3022
FT /*tag= a
FT /*product= "ZAP1 protein"
```

```
XX
XX WO200004173-A1.
XX
XX 27-JAN-2000.
XX
XX 12-JUL-1999; 99WO-EP04940.
XX
XX 17-JUL-1998; 98US-0118276.
XX
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
XX Babyichuk E, Kushnir S, De Block M;
XX
XX WPI; 2000-182436/16.
XX
XX P-PSDB; AAY68833.
```

Modulating cell death, growth and stress resistance in eukaryotes, specifically plants, used, e.g. to impart fungus or nematode resistance

Example 1; Page 79-84; 126pp; English.

The present sequence encodes the ZAP1 protein of Zea mays. This protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed cell death or apoptosis, and is a nuclear enzyme. The ZAP1 polynucleotide sequences can be used for modulation of programmed cell death in eukaryotic cells. The method is used, specifically in plants, to induce, or protect against, programmed cell death, depending on the extent to which PARP activity is reduced. Reducing expression of endogenous ZAP class PARP only is also used to modulate programmed cell death, to increase growth rate and to produce plant cells that are more tolerant of stress (cold, chemical treatments, pathogens, pests, drought, heat, etc., or during transformation). Particular applications are generation of plants that are resistant to fungi or nematodes; are male or female sterile; or have better seed-shatter properties. The methods are also used to improve growth of transformed plant cells (and derived calli or complete plants).

SQ Sequence 3211 BP; 968 A; 604 C; 813 G; 826 T; 0 other;

Query Match 92.2%; Score 2720; DB 21; Length 3211;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 2869; Conservative 22; Mismatches 16; Indels 45; Gaps 11;

Db 2204 gcagatccagcgcgcatactccatccctcagtgagggtccagcagcggt-----gtctcag 2258
 Qy 2076 ctatcaagcacgctgctgtgtrgagaaagcnaattgttgytgsgagcmatsyttttc 2135
 Db 2259 ggcagcagcactcagatcctggatctcctcaatcgcttttacacctcgtatcccccc 2318
 Qy 2136 actcttatccctctctatccatccatctattatcacggatgaggtatctcatattcaaa 2195
 Db 2319 gactttgggatgaagaagcctccgctccggaacaatgcaga----cagtggtcaggccaa 2374
 Qy 2196 ggcgaataatctgaagcctcgcagatattgaaattgctcaa-----agatagttgg 2249
 Db 2375 ggtggaatcttgcaaacctcgtgacatcgtgaggtgaggtgctcagctgctcaggggag 2434
 Qy 2250 ctctgatagcagcagtgatgaatctcttgatgataataatgaaacttcactgtgacat 2309
 Db 2435 gctgatgtagcagcaagatccctcgtatgcactatgagaagctcaaaactgacat 2494
 Qy 2310 caccgcgtggtcagcagatgtagaattacaaagttaattgagcagatctctcctcaaac 2369
 Db 2495 taaggtggttgacagagattctgaagaagccgagatcatcaggaagatgtttaagaacac 2554
 Qy 2370 acatgctctactacacagcagctgctgctggaactgaggaggttttttctcactgatog 2429
 Db 2555 tcatgcaac 2614
 Qy 2430 agatgagaaacttaataagctactcaagatataaaataatctgcatacaagaatctatt 2489
 Db 2615 tgaagcgaatccagcgttaacagccctt---aagcagctctcaacacccgaagattgct 2671
 Qy 2490 atggcaggttcaaggtgacgaattttggaatttttagtcaagggttaagaattgc 2549
 Db 2672 gtggcaggttcagaccacaaacttctggtgactctgtccaggtcttcgtagac 2731
 Qy 2550 acctctgagcagcctgttactgctgtatgttcggcaagcctctactcttgcagatct 2609
 Db 2732 ccgcctgaagcgcctgacaggtcacatgtttgtaagggatctatttcgctgacat 2791
 Qy 2610 agtaagcagagcgcacatactgttattgtagtaggaataatctctgtaggtttgact 2669
 Db 2792 ggtctcaagatgcacaaactgcatcactgctcaggagacccaaataggcttaactct 2851
 Qy 2670 tcttctgaggttctttaggagacatgtatgaactaaagaagccacgctccatggacaa 2729
 Db 2852 gttgggagaagttgcttcttggaacatgtatgaactgaagcagcttccacatcagcaa 2911
 Qy 2730 acctcaagaggaagcattcgaacaggttaggcaaaaacccgtgcactggagtcaga 2789
 Db 2912 gttaccacagggcagcagcagtgtaaaaggttgggcaaaaactaccctgactctcag- 2970
 Qy 2790 gttgtgaggtgagggatgctgtagtctcctgcggcagcagcgtgcccactcatcaat 2849
 Db 2971 --ctaactatgctggtggtgtagacgttctcttggaccgggattctatctggtgt 3028
 Qy 2850 taggagctctgaactcatgtataatgatactctcacaacacatcccaggtgaagat 2909
 Db 3029 gaatgacacctctactataaagcagatgataattgctatgattgctcaggtaaatct 3088
 Qy 2910 gcagttcttctgaaggtgcgtttccatcacaga 2944
 Db 3089 gaagtatctgctgaaactgaaattcaattttaaga 3123

RESULT 5

AAZ23799

ID AAZ23799 standard; cDNA; 3580 BP.

XX AAZ23799;

AC AAZ23799;

XX 18-JAN-2000 (first entry)

DE Human poly(ADP-ribose) polymerase cDNA.

XX

Gene therapy; poly(adenosine diphosphate-ribose) polymerase; treatment;
 PARP; antitumor; nuclear DNA repair; proliferating cell; DNA damage;
 protection; genomic instability; cancer; prevention; human; ds.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 140..3184
 /*tag= a
 /product= "poly(ADP-ribose) polymerase"

DE19808889-A1.

09-SEP-1999.

03-MAR-1998; 98DE-1008889.

03-MAR-1998; 98DE-1008889.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

Buerkle A, Meyer R;

WPI; 1999-509563/43.

P-PSDB; AAY33699.

New gene therapy vector expressing poly(adenosine
 diphosphate-ribose)-polymerase for treating or preventing tumors -

Claim 4; Fig 1; 12pp; German.

This invention describes a novel vector (A), suitable for gene therapy,
 which contains a DNA insert (I) that encodes a practically complete
 poly(adenosine diphosphate-ribose) polymerase (PARP) which has antitumor
 activity. PARP is a nuclear DNA repair enzyme (activated by strand
 breakage), that allows recovery of proliferating cells from the toxic
 effects of DNA damage (caused by alkylating or oxidizing agents or
 radiation), and protects cells against such damage or other causes
 of genomic instability. (A) are used for treatment of cancer and for
 cancer prevention in subjects at high risk (e.g. those with
 tumor-associated genetic defects). Tumor cells treated with (A) show an
 increased tendency to die when treated with radiation or chemotherapeutic
 agent, and in cells that survive this treatment they inhibit genomic
 instability, so should reduce the likelihood of further development of
 both malignant cells and resistance to chemotherapy. This sequence
 encodes the human poly(ADP-ribose) polymerase described in the invention.

Sequence 3580 BP; 986 A; 823 C; 996 G; 775 T; 0 other;

Query Match 7.5%; Score 221.8; DB 20; Length 3580;

Best Local Similarity 51.1%; Pred. NO. 1.1e-54;

Matches 692; Conservative 10; Mismatches 614; Indels 39; Gaps 7;

Qy 1602 ctactatgtactccagatcattgaacaggtgatgggtctggtcagctacgtatttcgttaa 1661

Db 1843 ctactacaagctcagctctctggagcagcaagaaacaggtattggtatttcaggtc 1902

Qy 1662 gtggggcaggggttggagtgagaaatttgaggggcaaaaactggaggagatgtcaaaaac 1721

Db 1903 ctggggcgtgtgggtacgtctcgttagcaacaactggaacagatgcgttccaagga 1962

Qy 1722 tgaggcaatcaaggaattcaaaagattttcttgagagactggaactggaactggaagc 1781

Db 1963 ggtatgcatgagcactctcagaattatagaagaaaacccgggaagcgttggcactc 2022

Qy 1782 ttgggaatgtaaaaccaattttcggagcagcctggagattttaccacttgatttga 1841

Db 2023 c-----aaaaatttcacgaagtatcccaaaaagtctaccctcgtgagattga 2070

Qy 1842 ttatggtttaaagaagcacc-----aaacggaagatatcagtgaaatgaaagttc 1895

Db 2071 ctatggccaggtgaagagcagtggaagagctgacagtaaatctctggcaccagtc 2130

Qy 1896 tcttgctcctcaattgctagaaactcatgaagatgcttttcaatgtggagacatatagagc 1955
Db 2131 gctccccaagcagctcaggaacctcatcaagatgcttttgatgtggaaagtatatgaaga 2190
Qy 1956 tgcctatgaggaattgaaataatataatgcagaatgcctcttctgggaagctaaagcmaggr 2015
Db 2191 agccatgttgagtagagacgacctcagaagaatgccttggggagagctgagcaaaag 2250
Qy 2016 aaatattgaaagagattgaaagcaataactkrgrtaacmratatttattgaaggacacg 2075
Db 2251 gcagatccagcgcgcatactccatccatcagtgaggtccagtagcggt-----gtctcag 2305
Qy 2076 ctatcaagcactgctgtttrgagaagcnaatgttgytgsagcmatsyitttttc 2135
Db 2306 ggcagcagcactcagatcctgctgctcctcaaatgcctttttacacccctgaccccccac 2365
Qy 2136 actcttatactctctctcctcctcctcattatcacgggagtgaggatgtattcattcaaa 2195
Db 2366 gactttggatgaagaagcctcctcctcctgaacaatgcaga----cagtgtagggccaa 2421
Qy 2196 ggcgaataatgctgaagctcgcagatgataatgaaattgcttcaa-----agatagttgg 2249
Db 2422 ggtggaatgcttgacaacctgctgacatcgaggtgaggtgaggtcagtgctcaggggag 2481
Qy 2250 ctctgtagcagcagtgatgatactctcttgatgataataatataatgaaacttcaactgtgacat 2309
Db 2482 gctgtagatagcagcaggaatccatcgatgctcaactatgagaagctcaaaactgacat 2541
Qy 2310 caccctgctgctcagcagtagtgaagattacaagtaattgagcagtgatctcctcaaacac 2369
Db 2542 taagtggtgacagagatcctgagaagcgcgagatcatcaggaagtgtttaagaacac 2601
Qy 2370 acatgctcctcactcaagagcagtgctgctggaactgaggaagtttttcaactgtagc 2429
Db 2602 tcatgacacacacaaatgcgtatgacttggaaatcgcgatactcttcaagatgagcgc 2661
Qy 2430 agatggagaacttaataagtaactgaatataataataataatctgcatacaagaatgctatt 2489
Db 2662 tgaagcgaatccagcagcttaaacgacctt--aagcagcttcaaacggaagattgct 2718
Qy 2490 atggcagcaggttcaaggttgacgaatttggggaattcttagtaaggcctcaagaattgc 2549
Db 2719 gtggcagcgggtccagagcaccacaaacttctgctgggactctgtccaggggtctcggatagc 2778
Qy 2550 acctcgaagcagcacttactgctgctatatttgcgcaagcctcactcttctgagatct 2609
Db 2779 cccgctgaagcgcgcgtgacgggtacatgcttggtaagggatctattctcgtgacat 2838
Qy 2610 agtaagcaagagcgcacaaactgttatgtgtagaggaataatcctctgtaggtttgtagct 2669
Db 2839 ggtctccaagagtccaactactgctcactgctcagggagacccaataggcttaactcct 2898
Qy 2670 tcttctgaggttctttaggagacatgataaataaagaagcagcagcctcactgagcaaa 2729
Db 2899 gttgggagaagttgctcccttggaaacatgatgaactgaagcagcgtctcactacatcagca 2958
Qy 2730 acctcaagagaggaagcattcgcacaaaggtattaggaacaaacgtgcccactggagtcaga 2789
Db 2959 gttaccaagggcagcagcaggtgtaaaaggttttgggcaaaactaacccctgactcctcag- 3017
Qy 2790 gtttggaggtgagggatgtagtgcctcctcctgctgagcagcgggtgcccactcatcaat 2849
Db 3018 --ctaactattgcttgaggtgtagacgttctccttctggaacgggatttcaactcgtgtg 3075
Qy 2850 tagggagctgaactcagtaacatgagtagacatcgtctcactcaacacacatccaggtgaagat 2909
Db 3076 gaatgacacctctcactataacagacattgctatgctatgattgctcaggtgaatct 3135
Qy 2910 gcagttcttctgcaaggtgcgttttccatcacaga 2944
Db 3136 gaagtattcgtgaaactgaaattccaaattttaaga 3170

RESULT 6

AAF63954
ID AAF63954 standard; DNA; 3045 BP.

XX AAF63954;

XX 05-APR-2001 (first entry)

XX Human tankyrase2 related coding sequence SEQ ID NO: 136.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX inflammatory disorder; ds.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI: 2001-102896/11.

XX P-PSDB; AAB66296.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders -

XX Example 2; Page 203-207; 242pp; English.

XX The present invention provides the protein and coding sequence for the
XX human tankyrase2 protein. This is found in two different versions,
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX polyADP-ribosylation activity and is involved in the modification of
XX TRF1, which is a telomere-specific binding protein. The regulation of
XX telomere length, in which TRF1 has a role, is linked to ageing and
XX cancer. The sequences are useful in the treatment of cancers and
XX inflammatory disorders.

XX Sequence 3045 BP; 863 A; 710 C; 851 G; 621 T; 0 other;

Query Match 7.5%; Score 219.8; DB 22; Length 3045;

Best Local Similarity 50.9%; Pred. No. 3.8e-54;

Matches 690; Conservative 11; Mismatches 615; Indels 39; Gaps 7;

Qy 1602 ctactatgtactcagatcattgaacagatgatgggtcgtgactgctagctatctgtaa 1661
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Db 1704 ctactacaagctgcagctctctggagcagcaaggaacagglattggatattcaggtc 1763
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Qy 1662 gtggggacgggttggagtgagaaattggaggcacaacaaactggagagatgtcaaac 1721
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Db 1764 ctggggccgtgtgtgacgtgacgtgtagcaacaaactggaacagatgccgtccaagga 1823
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Qy 1722 tggagcaatcaagaattcaaaagattattcttgagaagactggaaactcatgggaagc 1781
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Db 1824 ggaagcattgagcagttcattgaattatgaagaaaaaacgggaacgttggcactc 1883
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Qy 1782 ttgggaatgaaacaaattttcggagcagcctggagatttttaccactttagttga 1841
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Db 1884 c-----aaatttcacagtgatccccaaagttttacccctggagattga 1931
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Qy 1842 ttaagtggttaagaagacacca-----aacggaaagatatcagtgaaatgaaagttc 1895
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||Db 1932 ctatggccaggatgaagagcagtgagaagctcacagtaaatcctctggcaccagtc 1991
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy	1896	tctgtcctcoattgctagaacctaagatgcttttcaatgtggagacatatagac	1951
Db	1992	gtccccaagcccagtctcaggacctcaagatgatcttgatgtggaagtatgaaga	2051
Qy	1956	tgcatagtgaatttgaawtaatatgtcagaaatgctcttgggaagctaagcmaggr	2015
Db	2052	agccatggttggagtagatcgaccctcagaagatgcccttgggaagc/gagacaag	2111
Qy	2016	aaattattgdraagagattgaagcatctaactkrgrtaomgratttatttgaagcacccg	2075
Db	2112	gcagatccagggcgatactccatcctcagtgaggtccagcaggggt----gtctcag	2166
Qy	2076	ctnatcaagcaactggctgttgrgaaaaactnaattgttytgs9agomatssytttttc	2135
Db	2167	ggcagcagcactcctcagatcctgtgattctctcaatcgcttttacocctgatccccac	2226
Qy	2136	actctatccctcttatcatcctcataatacacgggatgagatgatttcatttcaaa	2195
Db	2227	gaetttggagaaagacgtccgtcctgyaacaatcgaga---cagtgtcgagggccaa	2282
Qy	2196	ggogaaaatcgttgaagctctcaggagatttgaaattgtctcaa-----agatagttgg	2249
Db	2283	ggtggaatcgttgacaacctgctggacatcgaggtgacctacagttotgctcagggagg	2342
Qy	2250	cttcgatagcgacagtgaatactcttgatgaataataataatgaaacttcaactgtgacat	2309
Db	2343	gtctgatagtagcagaaggtacccatcgatcgatcaactgatgagaagctcaaaactgcac	2402
Qy	2310	cacccgcgtggtcacgatalgtgaagattacaagttaaattgagcagtatctctcaaac	2369
Db	2403	taagg/tggttgcagagagattctgagaagcgcgagatcatcaggaagtgtttaagaacac	2462
Qy	2370	acatgctcctactcacaaagacgtgtcgctggaaactggaggaagtgttttcaacttgatcg	2429
Db	2463	tcattgaaaccaacacagtcgctatgacttggagtgatcgatctcttaagatagcgg	2522
Qy	2430	agatggagaacttaaataagtctcaagataataaaaataactctgatacaacagatgctatt	2489
Db	2523	tgaaggcgaatgccagcgttacaaagccctttaagcag--cttcataacccgaagattgct	2579
Qy	2490	atggcagcgttcaagctgaacgaatttgttgggaattcttatgtcaagggctaagaattgc	2549
Db	2580	gtggcagcgggtccaggaccacaacttctgctggatcctgttcccaggggtcttcoggatgc	2639
Qy	2550	acctcctgaggcacctgttactggctatgtctggcaaaagcctctactttgcacatct	2609
Db	2640	ccgcgtgaagggccgtgacaggtacatgttggtaaaagggatcattctcgtgacat	2699
Qy	2610	agtaacgaagagocacaaactgttattgttgatggaataaactcctgtaggtttgatgct	2669
Db	2700	ggctccaaagtgccaactactacatacgtctcaggggagaaccaataaggcttaactcct	2759
Qy	2670	tccttctgaggtgctttaggagacatgtatgaactaaagaagaacccagctccatggacaa	2729
Db	2760	gttggagagagttgccttggaaacatgtatgaactgaagcaagctgcacatatacgcag	2819
Qy	2730	acctcgaagaggaagcatttcgaacaggattaggcaaaacccgtgccactggagtcaga	2789
Db	2820	gttaaccaagggccaagcacagtgtaaaaggtttggggcaaaactaccctgatccttcag-	2878
Qy	2790	gttttgaagtggagggatgatgtgtagtttccctgcggcaagccggtgccatcatcaat	2849
Db	2879	--ctaacaattagcttggatgggtgtagcgttccctcttgggaccgggatttcacttggtgt	2936
Qy	2850	taggagctctgaactcatgtacaatgagtacatcgctctacaacacatcccaggtgaagat	2909
Db	2937	gatagacacctctactatataaocagatgcatgtctatgatatgtctcaggtaaatct	2996
Qy	2910	gcagttcttctgctgaaggtgcgttttccatcacaga	2944
Db	2997	gaagtatctgctgaactgaataatcaattttaaga	3031

RESULT	7
AAC85341	
ID	AAC85341 standard; cDNA; 3200 BP.
XX	
AC	AAC85341;
XX	
DT	29-MAR-2001 (first entry)
XX	
DE	Fusion protein PARPIA/PAR2B CDNA.

Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS; inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia; infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis; gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock; endotoxic shock; gram negative sepsis; gram positive sepsis; trauma; toxic shock syndrome; multiple organ injury syndrome; vasculitis; hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy; eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis; chronic obstructive pulmonary disease; silicosis; reperfusion injury; pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium; bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain; scar tissue formation; atherosclerosis; systemic lupus erythematosus; autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome; graft versus host disease; allograft rejection; cystic fibrosis; chronic glomerulonephritis; inflammatory bowel disease; Crohn's disease; ulcerative colitis; necrotizing enterocolitis; inflammatory dermatosis; contact dermatitis; atopic dermatitis; psoriasis; urticaria; fever; myalgia; meningitis; encephalitis; Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia; hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity; leukocyte dyscrasia; thermal injury; cytokine-induced toxicity; expressed sequence tag; EST; RACE; PCR; amplify; primer; polymerase chain reaction; ds.

Synthetic.

```
key Location/Qualifiers
CDS 1..3200
    /*tag= a
    /product= "PARP1A/PARP2B"
```

WO200077179-A2.

21-DEC-2000.

16-JUN-2000: 2000WO-US16629.

16-JUN-1999: 99US-0139543

(TCOS-) TCOS CORP.

Christenson E, Demaggio AJ, Goldman PS, McElligott DL:

WPI: 2001-025335/03.

P-PSDB; AAB4/U3Z.

new human poly(ADP-ribose) polymerase for treating inflammatory, neurological, cardiovascular, or neoplastic tissue growth disorders, such as, arthritis, encephalitis, myocardial ischemia, and leukocyte metastasis -

Example 3; Page 116-117; 129pp; English.

This sequence encodes the fusion protein hPARP1A/hPARP2B. This protein contains amino acids 1-662 of hPARP1 fused upstream of amino acids 230-583 of hPARP2. This sequence was amplified using the primer sequences given in AAC85321-40 and AAC85342-51. The fusion protein coding sequence was cloned in a baculovirus expression vector for the expression of the fusion protein. The protein of the invention, hPARP2, causes the covalent addition of polymers of ADP-ribose to protein targets. hPARP2 activity is induced in many instances of oxidative stress or during inflammation where there is direct damage to the DNA. hPARP2 may be used to identify

CC antagonists which may be used to treat a human having a disorder
 CC mediated by PARP2 activity, such as, inflammatory, neurological,
 CC cardiovascular, or neoplastic tissue growth disorders. hPARP2 and
 CC antibodies to it, can also be used to diagnose these conditions.

XX Sequence 3200 BP; 923 A; 745 C; 870 G; 662 T; 0 other;

Query Match 7.0%; Score 205.2; DB 22; Length 3200;
 Best Local Similarity 49.9%; Pred. No. 8.2e-50;
 Matches 675; Conservative 9; Mismatches 634; Indels 36; Gaps 6;

Qy 1602 ctactatgtactcagatcattgaacagatgattggtctgtgagtgctacgtatttcttaa 1661
 Db 1812 ctactacaagctgcagctcttgaggagcagcaagaaacagaggtatggatcattcaggtc 1871
 Qy 1662 ttggggacgggttgagtgagaaatggaggggcacaacaaactggagagatgtccaaaac 1721
 Db 1872 ctggggccggttggtgtacgggtgacgtagcaacaaactggaacagatgcccgtccaaaga 1931
 Qy 1722 tgaggcaatcaaggaaattcaaaagattattcttgagaagactggaaactcatgggaagc 1781
 Db 1932 ggatgccattgagcattcatgaaattatgaagaaacacgggaacgcttgccactc 1991
 Qy 1782 ttgggaatgtaaaacaaattttcgaag-----cagcctgggagattttaccacattga 1835
 Db 1992 caaaaatttcacgaagtatcccaaaaagtctacccctggagattgactatgcccagga 2051
 Qy 1836 tgttgattatggtttaagaaagcaccacaaacggaaagatatcagtgaaatgaaaagtcc 1895
 Db 2052 tgaagaggcagtgaaagactgacagttaactctggcaccacacgcttccagagtcaca 2111
 Qy 1896 tctgtctctcaattgctgagactcatgaagatgttttcaatgtggagacatatagagc 1955
 Db 2112 gctagatcttcgggtacagaggttaataaagttgtatctgtcaatgtccagggccatggaaga 2171
 Qy 1956 tgcctatgatgaattgaaatgaatattgtcagaataatgctctgtggagcgttaagcmaggr 2015
 Db 2172 aatgatgatggaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 2231
 Qy 2016 aatattgagragagattgaaagcattaaactrrgtacmrrattttatttgaaggacacgc 2075
 Db 2232 gcaaatcaagcaggttaccagctctcttaagaagattgaggtattgtattcgggtggcca 2291
 Qy 2076 ctatcaacagcactggtttgttgagaagcnaattgttgytgsagcmattsytttttc 2135
 Db 2292 gcatgacgagctcatggaagcagtcgaatgaattcta-----caccagagattcc 2342
 Qy 2136 actcttatccctcttattctctcattattatatacgggagtgaggtatttctatttcaaa 2195
 Db 2343 gcatgactttgactccgtactcctcactaatccggcacacagaagaaactgtcagaanaa 2402
 Qy 2196 ggcgaaaatgcttgaagctctgcagatatgtgaaattgttccaaagatgttggttcga 2255
 Db 2403 aatacaattactagagccttggagacattgaaattgtattaaagctggtgaaacacaga 2462
 Qy 2256 tagcgacag---tgatgaatctcttgatgataaataatataaacttcaactgtgacatcac 2312
 Db 2463 gctacaagcccgacacacccattggaccacactatagaaacactacattgtgccttgcg 2522
 Qy 2313 cccgctggtcagcagatagtgaaattacaagtttaattgagcagatattctctcaacacaca 2372
 Db 2523 ccccttgaccatgaaagtacaggttcaagtgatgatttccctcagttaccataactaccaca 2582
 Qy 2373 tgcctctactcaagaagactggtgcgtggaaactggagagagatttttttcttctgactcaga 2432
 Db 2583 tgcctccacacacagcactataccatgacactgtgtgatttcttgaagtggagaagga 2642
 Qy 2433 tgggagaacttaataagttactcaagataataaaaaaataactctgatacaacagatgctattat 2492
 Db 2643 tgg-----tgagaaagaagccttcagagagaccttcataacagagatgctctcatg 2693
 Qy 2493 gcaagggttcaaggttgacgaattttgtgggaattcttagtcaagggtcaagaattgcacc 2552

Db 2694 gcatggtccaggatgagtaactgggtgggaatcttgagccatgggtctcgaaatgcccc 2753
 Qy 2553 tctgaggcacctgttactggtctatatgttcgcaaaagcctctacttttgcagatctagt 2612
 Db 2754 acctgaagctcccatcacaggtttacatgttttgggaaagaaatctactttgtgcagatgc 2813
 Qy 2613 aagcaagagcgcacaaatactgttatgtgagtaggaataatcctgttaggtttgactcttct 2672
 Db 2814 ttccaagagtccaattactgctttgtccctcgtcctaaagaatacaggagactgctctct 2873
 Qy 2673 tctgagggtgttctttaggagacatgtatgaactaaagaagccagcctccatggacaaaac 2732
 Db 2874 atcagaggtagctctcaggtcagtgtaataagactactagagggccaatcctaaagccgaag 2933
 Qy 2733 t-----ccaagaggaagcattcgaccagggtattagcacaacccgtgccaactggagtcaga 2789
 Db 2934 atgtctcaaggttaacatagcaccagggcgtggcagagatggtcccgagttctgcca 2993
 Qy 2790 gtttg-----tgaagtgaggagatgctgtagttccctgcggcgaagccggtgccatc 2843
 Db 2994 ctctgcacccctgaatgggagtagcagtgccatttagdaccagcaagtgcacacaggaattct 3053
 Qy 2844 atcaattagagagctgaactcatgtacaatgagtagcatcgtctacacacacatcccagggt 2903
 Db 3054 gaatccagatggtttataccctcaactcaactacaatgaatatattgtatataaccccaaccaggt 3113
 Qy 2904 gaagatcagttcttctgctgaaggtgcgtttccat 2937
 Db 3114 ccgtatcggtacaccttttaaaaggttcagtttaat 3147

RESULT

AAZ60617
 ID AAZ60617 standard; DNA; 2147 BP.

AC AAZ60617;

DT 16-MAY-2000 (first entry)

DE DNA encoding the poly(ADP-ribose) polymerase NAP protein.

KW NAP: poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;
 KW Programmed cell death; apoptosis; growth rate; stress; cold; pathogen;
 KW pest; drought; heat; fungi; nematode; seed-shatter; ss.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers

FT CDS 129..2042

FT /*tag= a

FT /product= "NAP protein"

XX WO200004173-A1.

PN 27-JAN-2000.

PF 12-JUL-1999; 99WO-EP04940.

PR 17-JUL-1998; 98US-0118276.

XX (PLBZ) PLANT GENETIC SYSTEMS NV.

XX Babyichuk E, Kushnir S, De Block M;

XX WPI; 2000-182436/16.

XX P-PSDB; AAY68835.

XX Modulating cell death, growth and stress resistance in eukaryotes,
 XX specifically plants, used, e.g. to impart fungus or nematode resistance

XX Disclosure; Page 95-99; 126pp; English.

XX This sequence represents cDNA encoding a novel human poly(ADP-ribose)
CC synthetase, shbPARS2. The invention also relates to fragments,
CC variants and sequences with at least 95% identity to the shbPARS2
CC protein or nucleotide sequence; expression systems and host cells
CC comprising an shbPARS2 nucleic acid sequence; the recombinant expressions
CC of shbPARS2; and an antibody specific for shbPARS2. shbPARS2 proteins
CC and nucleotides are useful as vaccines for inducing an immunological
CC response in a mammal. The shbPARS2 protein is useful for identifying
CC compounds which inhibit or stimulate its activity or expression level
CC Such agonists and antagonists of shbPARS2 are useful for treating human
CC diseases including ischemia and ischaemic tissue injury (e.g., cerebral,
CC and cardiac ischemia, myocardial infarction, stroke), inflammation,
CC and autoimmune disease (e.g. diabetes, multiple sclerosis) and
CC neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's
CC disease). shbPARS2 nucleic acids are useful as diagnostic reagents for

detecting mutations in the associated gene; as hybridisation probes to isolate full-length sbPARS2 cDNAs and sbPARS2 genomic clones; and for chromosome localisation studies. The sbPARS2 protein is also useful as an immunogen to produce antibodies for therapeutic use. sbPARS2 proteins, nucleotides and antibodies are also useful in screening methods for detecting the effect of added compounds on the production of mRNA and protein in cells.

SQ Sequence 1566 BP; 490 A; 328 C; 381 G; 367 T; 0 other;

Query Match 6.5%; Score 191.8; DB 22; Length 1566;
Best Local Similarity 51.0%; Pred. No. 4.8e-46;
Matches 570; Conservative 9; Mismatches 508; Indels 30; Gaps 5;

5:

QY	1833	tgatgttgattatggtgtttaagaaagaccacaaacggaagaaatcatcagtgaatgaagaag	1899
Db	450	tcagatgagagagaaacaaagaaaggaatctcttaaatctcccttgaagccagatc	509
QY	1893	ttctcttgctctcaattgctagaacctgaagatgottttcaatgttgagacatatag	1952
Db	510	acagctagatcttcgggtacagaggttaataaagttgatctgtaatgttcaggccatgga	569
QY	1953	agctgctatgatggaatttgaaawtaatatgtcagaaatgcctcttgggaagctaaagcma	2012
Db	570	agaaatgatgagaaatgaagtataatacacaagaaagcccccacttgggaagctgacagt	629
QY	2013	ggaaaaataggaaagatttgaagcatttaactkrgtaomgratttatttgaagaca	2072
Db	630	ggcaaaatcaaggaggttacagctctcttaagaagatggagatggtatctcggtctgg	689
QY	2073	ccgctnataagcaactggctttrtgagaagagctnaattgttgytgsagcmatsytttt	2132
Db	690	ccagcatggagagctctcatgaagcatgcaatgaattctta-----caccaggt	740
QY	2133	ttcaactttatcccttattcatctcatattatcaggaatgagatgatttcatttc	2192
Db	741	tcgcgatgactttggactcgttactctccactaatcggcacagaaagaaactgctcaga	800
QY	2193	aaagcgcgaataatgcttgaagctctgcagatattgaaattgcttcaagaatagttggctt	2252
Db	801	aaaaatcacaaatctactagaggtcttgggagacattgaaattgctattaaactgttgaaaac	860
QY	2253	cgatgcgcagctgatatctc-----ttgatgataaataatataaaacttcaactgtgacat	2309
Db	861	agagctacaaagcccgacaacccacttggaccaacatatagaacctacatgttgcctt	920
QY	2310	caccgcgtgcctcacgatatgtgaagattacaagtttaattgagcagtatctcctcaaac	2369
Db	921	gcgcccccttgaccatgaagattgatgttgcgaatttcaagtttccagctactcataactac	980
QY	2370	acatgctctactcacaaagactggtgcctggaaactggaggaagtttttcaacttgcacg	2429
Db	981	ccatgctccacacacagcgactataccatgaacctgcttggattgttttgaagtggagaa	1040
QY	2430	agatggagaacttaataagttactccaagataaaaaataatctgcatacaacagatgctatt	2489
Db	1041	ggatgg-----tgaaaaagagccttcagagagggaccttcaataacagagtgcttct	1091
QY	2490	atggcacggttcaaggtttgacgaatttttgggaattcttagtcaagggctaaagaattgc	2549
Db	1092	atggcatggttccagatgtagtaactgggtcgggaaacttgagccatgggctcgaattgc	1151
QY	2550	acctctgagcgacctgttactggcttatattgttcggcaagggctctactttgcagatct	2609
Db	1152	cccactgaagctcccactcaggttacatgtttggaaaagaaactacttcttgcagat	1211
QY	2610	agtaagcaagcgccacaaatcactgttatgttgataggaataactcctgtaggtttgcagct	2669
Db	1212	gtcttccaagagtcccaattactgttttgccttgccttaagaatacaggaactgcgtct	1271
QY	2670	ttcttctgaggttgcctttaaagagacatgtatgaactaaagaagccagctccatggcaca	2729

Db	1272	cttatcatgaggtagctctaggtcagtgtaatgaactactagagccaaatctctaagccga	1333
QY	2730	---accctcaagagggaagcattcgaacaaaggattagcgcaaaacgctgccactgagtc	2786
Db	1332	aggatgctcaagtgataacatgacaccaaaggcgtggcaaatggtctccagttctctgc	1391
QY	2787	agagtttg-----tgaagtggaggatgatgtcgttagttccctgcgcaagcgggtgc	2840
Db	1392	ccacttcgcacccctggaatggagtagcagtgccattaggaccagcaagtgacacaggaat	1451
QY	2841	atcataaattaggagctctgaactcatgtacaatgaatgacatcgtctacaacaacatccca	2900
Db	1452	tctgaatccagatggtataccctcaactcaaatgaaatattgtatataaaccacaacca	1511
QY	2901	ggtagaatgcagttcttctgtgaaggtgcgtttccat	2937
Db	1512	ggtcgtagtgggtacaccttttaagggtcagtttaat	1548
RESULT 11			
AAC85303			
ID	AAC85303	standard; cDNA; 1814 BP.	
XX	AAC		XX
XX	AAC		XX
XX	29-MAR-2001	(first entry)	
XX	hpap2	cDNA.	
XX	Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;		
KW	inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;		
KW	infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;		
KW	gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;		
KW	endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;		
KW	toxic shock syndrome; multiple organ injury syndrome; vasculitis;		
KW	hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;		
KW	eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;		
KW	chronic obstructive pulmonary disease; silicosis; reperfusion injury;		
KW	pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;		
KW	bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;		
KW	scar tissue formation; atherosclerosis; systemic lupus erythematosus;		
KW	autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;		
KW	graft versus host disease; allograft rejection; cystic fibrosis;		
KW	chronic glomerulonephritis; inflammatory bowel disease;		
KW	Crohn's disease; ulcerative colitis; necrotizing enterocolitis;		
KW	inflammatory dermatosis; contact dermatitis; atopic dermatitis;		
KW	psoriasis; articularia; fever; myalgia; meningitis; encephalitis;		
KW	Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;		
KW	hypovolemic shock; type 1 diabetes mellitus; hypersensitivity;		
KW	leukocyte dyscrasia; thermal injury; cytokine-induced toxicity; ds.		
XX	Homo sapiens.		
Key	Location/Qualifiers		
CD	63..1814		
TT	/*tag= a		
TT	/product= "hPARP2"		
W	WO200077179-A2.		
X	21-DEC-2000.		
D	16-JUN-2000; 2000WO-US16629.		
F	16-JUN-1999; 99US-0139543.		
R	(ICOS-) ICOS CORP.		
A	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;		
I	WPI; 2001-025335/03.		
X	P-PSDB; AAB47029.		

Qy	1833	tgatgtgattatggctttaaagaagcccaaaacggaagaatacatcagtgaatgaaaag	1899
Db	698	tcaggatgagaggaacaagaagagaaactcttaaatctccccttgaaaccagatgc	757
Qy	1893	ttctctgtccccaattgctagaactcatgaagatgctttctaattgtagacatatag	1952
Db	758	acagctagatcttcgggtcacagaggttaataaagttgatctgaatgtcagggccatgga	817
Qy	1953	agctgctatgatgaattgaaawtaatatgcagaaatgcctcttggaagctaaagcma	2012
Db	818	agaaatgatgatggaatatgaagcataataaccaagaagccccaccttgggaagctgcacagt	877
Qy	2013	ggraaatattgagraagagatttgaagcattaactkrgttacmcratttatatttgaaggaca	2072
Db	878	ggcacaaatacgaagcgaggttacagctctcttaagaagattgaggattgtattcgggctgg	937
Qy	2073	ccgctnatcaagcacactggcttgtttgagaaaagcnaattgttgytgsagcmatacssyttt	2132
Db	938	ccagcatgagcagactctcatggaagcatgcaatgaattcta-----caccaggat	988
Qy	2133	ttcaactttaacccttctattcatcctcaattatacogggatgaggatgtattcatattc	2192
Db	989	tcqcatgaactttggactccgtactctccactaactcgaacacagaaggaactgtcaga	1048

Fri Mar 8 08:31:51 2002

us-09-236-995d-1.rng

Page 20

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 14:33:30 ; Search time 122.78 Seconds
(without alignments)
5439.678 Million cell updates/sec

Title: US-09-236-995D-1
Perfect score: 2949
Sequence: 1 atggcgccgcgcaagc.....gtttccatcacagaggttag 2949

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	219.8	7.5	3747	1 US-08-044-618-5	Sequence 5, Appli
2	160	5.4	5345	1 US-08-044-618-7	Sequence 7, Appli
3	158.4	5.4	2682	1 US-08-044-618-3	Sequence 3, Appli
4	131	4.4	1592	1 US-08-044-618-4	Sequence 4, Appli
5	125	4.2	595	4 US-09-328-111-216	Sequence 216, App
6	61.2	2.1	7218	1 US-08-232-463-14	Sequence 14, Appli
7	39.2	1.3	2313	2 US-08-892-880-1	Sequence 1, Appli
8	35.4	1.2	6060	5 PCT-US96-09430-7	Sequence 7, Appli
9	35.2	1.2	336	1 US-08-331-398A-22	Sequence 22, Appli
10	35.2	1.2	336	1 US-08-331-397B-22	Sequence 22, Appli
11	35.2	1.2	336	4 US-08-759-804A-22	Sequence 22, Appli
12	35.2	1.2	336	4 US-09-227-693-22	Sequence 22, Appli
13	35.2	1.2	375	1 US-08-331-398A-60	Sequence 60, Appli
14	35.2	1.2	375	2 US-08-331-397B-60	Sequence 60, Appli
15	35.2	1.2	375	2 US-08-759-804A-59	Sequence 59, Appli
16	35	1.2	722	4 US-09-328-111-681	Sequence 681, App
17	35	1.2	1315	3 US-08-908-332-3	Sequence 3, Appli
18	34.8	1.2	342	1 US-08-285-936-3	Sequence 3, Appli
19	34.8	1.2	342	1 US-08-487-860-3	Sequence 3, Appli
20	34.6	1.2	289	4 US-09-007-005-17	Sequence 17, Appli
21	34.6	1.2	289	4 US-09-244-796-17	Sequence 17, Appli
22	34.4	1.2	3180	2 US-08-500-857A-9	Sequence 9, Appli
23	34.2	1.2	8931	3 US-09-051-019-1	Sequence 1, Appli
24	34	1.2	1204	1 US-08-181-271A-21	Sequence 21, Appli
25	34	1.2	1204	1 US-08-449-315-21	Sequence 21, Appli
26	34	1.2	1204	1 US-08-444-803-21	Sequence 21, Appli
27	34	1.2	1204	1 US-08-449-043-21	Sequence 21, Appli

28	34	1.2	1204	1 US-08-456-265A-21	Sequence 21, Appli
29	34	1.2	1204	1 US-08-455-416-21	Sequence 21, Appli
30	34	1.2	1204	1 US-08-455-244-21	Sequence 21, Appli
31	34	1.2	1204	1 US-08-454-876-21	Sequence 21, Appli
32	34	1.2	1204	2 US-08-457-364-21	Sequence 21, Appli
33	34	1.2	1204	2 US-08-456-262-21	Sequence 21, Appli
34	34	1.2	1204	2 US-08-456-240-21	Sequence 21, Appli
35	34	1.2	1204	2 US-08-455-736-21	Sequence 21, Appli
36	34	1.2	1204	4 US-08-971-217-21	Sequence 21, Appli
37	34	1.2	1204	4 US-09-350-600-21	Sequence 21, Appli
38	34	1.2	1399	3 US-08-335-469-1	Sequence 1, Appli
39	34	1.2	1522	2 US-08-481-658B-1	Sequence 1, Appli
40	34	1.2	1522	2 US-08-477-504A-1	Sequence 1, Appli
41	34	1.2	1522	2 US-08-486-756A-1	Sequence 1, Appli
42	34	1.2	1522	2 US-08-485-862B-1	Sequence 1, Appli
43	34	1.2	1522	3 US-08-787-739-1	Sequence 1, Appli
44	34	1.2	1522	3 US-08-487-077A-1	Sequence 1, Appli
45	34	1.2	1522	3 US-08-485-863A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-044-618-5
; Sequence 5, Application US/08044618
; Patent No. 5449605
; GENERAL INFORMATION:
; APPLICANT: SMULSON, MARK
; TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO
; TITLE OF INVENTION: CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH
; TITLE OF INVENTION: POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,618
; FILING DATE: 19930406
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/257,696
; FILING DATE: 14-OCT-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAMUEL L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0654.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3747 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-044-618-5

Query Match 7.5%; Score 219.8; DB 1;
Best Local Similarity 50.9%; Pred. No. 2.le-55;
Matches 690; Conservative 11; Mismatches 615; Indels 39; Gaps 7;

[illegible]

```

RESULT      5
US-09-328-111-216/c
; Sequence 216, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: Fastseq for Windows Version 3.0
; SEQ ID NO 216
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(595)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-216

```

21	Db	994	CAGAGGACTCCATCGATGTCAACTATGAGAAGCTCAAAACTGACATTAAGSTGGTTGA	1053
22	Qy	2322	tcacgtagtagaagtatacaagttaattgacgtagtatctctcaacacacatgctctctac	2381
23	Db	1054	CAGAGATTCTGAAGAGCTGAGATCATCAGGAAGTATGTTAAGAACACTCATGCAACCA	1113
24	Qy	2382	tcacaaggactggtcgctggaactggaggaagtttttcaacttgatcgaga---tgagaga	2438
25	Db	1114	CCACACACGATG--CATATGACTTGGAACTCATTCATAGCTTTTAAGATAGAGTGTGAAGA	1171
26	Qy	2439	acttaataagtagtactcaagatataaaaataatctcataacaagatgctattattggaacgg	2498
27	Db	1172	GGAGTCCACGACACTACAGGCCCTTTAAGCAGCTTCATTAACCTGAAGTTGCTGTGGCATGG	1231
28	Qy	2499	ttaagggttgacgaatttttgggaattcttagtcaaggcgctaagaattgcaactcctga	2558
29	Db	1232	GTCCAGGACCACTTTTGCTGGGATCCTTCGCTCCCTTTGGATAGCCCTGCCTGA	1291
30	Qy	2559	ggcacctgttactgctctatatgttgcgcaaggcctctactcttgcagatctagtaagcaa	2618
31	Db	1292	AGCACTGTGATGGGCTACATGTTTGGTAAAGTAGTATCTTCGCTGATCTTCTGCCAA	1351

	Query Match	4.2%;	Score 125;	DB 4;	Length 595;
	Best Local Similarity	60.3%;	Pred. No. 1.5e-27;		
	Matches 225;	Conservative 0;	Mismatches 145;	Indels 3;	Gaps 1;
OY	2408	aggaagtgttttttcacttgatcgatgaggaacttaataagtactcaagataaaaaata	2467		
Db	410	ATGACCTTGCTGGATTGTGTTGAAGTGAGAAGGATGGTGAGAAAGAGCCTTCACAGAGG	351		
OY	2458	atctgcataacaagaatgctctattatgcaocggttcaagggttgacgaattttgtgggaattc	2527		
Db	350	ACCTTCATAACAGAGATGCTTCTATGGCATGGTTCACAGGATGATTAACCTGGGTGGCAATCT	291		

2528	ttagtcaagggctaaagattgcacctctctgaggcacctgttactggtctatatatttcggca	2587
QY		
290	tgaggcattggccttcgaattgccccacctgaagctgcccatcacaggttacattgttggga	231
Db		
2588	aaggcctctactttgcagatctagtaagcaagagcgcacaactactgttatgtggatagga	2647
QY		
230	aaggaaatctacttttgcgtacatgcttccaaagatgcgaattactgcgttgcctctcgcc	171
Db		
2648	ataatccctgtaggtttgatgcctttctctgaggtgtcttaggagacatgatgaactaa	2707
QY		
170	taagaattacaggactgcgtctttatcacaggtagctctagctgcagtgttaattgaactac	111
Db		


```
;
; NAME/KEY: 1..375
; LOCATION: 1..375
; OTHER INFORMATION: /note= "Mouse monoclonal antibody B5 Fv
; OTHER INFORMATION: Light chain region"
US-08-759-804A-59
```

```
Query Match 1.2%; Score 35.2; DB 2; Length 375;
Best Local Similarity 53.2%; Pred. No. 0.86; Indels 0; Gaps 0;
Matches 67; Conservative 3; Mismatches 56; Indels 0; Gaps 0;

QY 1457 ctgtcaactgttaagtttaagggccgaagtgtgttttcattgaagttgcaaga 1516
Db 35 CTGTCAGTCTTGGAGATCAAGCCCTCTATTCTTGTAGATCTAGTCAGAGCATTGTACATA 94

QY 1517 atactgtcacattctttragrattggaaagcatatatacaatgcamecttaaacatgttc 1576
Db 95 GTAATGGAACACCTATTATTAGATGTACTCTCAGAAACCCAGGCCAGTCTCCAAAGCTCC 154

QY 1577 tgacct 1582
Db 155 TGATCT 160
```

Search completed: March 7, 2002, 16:49:07
Job time: 8137 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2002, 14:28:50 ; Search time 2934.5 Seconds
(without alignments)
10798.876 Million cell updates/sec

Title: US-09-236-995D-1
Perfect score: 2949
Sequence: 1 atggcggcgccgcaaaagc.....gttccatcacaaagtag 2949

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_estl:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	579.8	19.7	653	10	AW066305
2	484.6	16.4	714	10	AV833309
3	338.2	11.5	607	11	BF255013
4	312	10.6	536	10	BE419432
5	301.8	10.2	427	10	BE516130
6	299	10.1	661	10	AV834168
7	269.2	9.1	369	10	BE420229
8	223.8	7.6	419	11	BE941860
9	221	7.5	674	11	BF459972
10	189	6.4	502	11	BG042229
11	186.2	6.3	759	13	AQ577141
12	181	6.1	669	10	AA401836

13	180.8	6.1	861	11	BG280821
14	174.6	5.9	916	11	BI093436
15	170.8	5.8	990	11	BG031594
16	163.2	5.5	728	11	AI455188
17	162.2	5.5	632	11	BG429402
18	161.4	5.5	661	10	AV833893
19	159.2	5.4	834	11	BI091452
20	159	5.4	678	11	BF101770
21	155.8	5.3	442	11	D42790
22	155.8	5.3	617	10	AA212857
23	155.4	5.3	716	11	BI217470
24	153.8	5.2	587	11	BE911127
25	153.2	5.2	560	11	BF042005
26	152.6	5.2	771	11	BG751755
27	152.2	5.2	607	11	BG816404
28	152	5.2	1113	11	BF976506
29	151.2	5.1	499	10	BE488132
30	151.2	5.1	571	10	BE304535
31	150.6	5.1	725	11	BI079902
32	148.2	5.0	579	10	AA397988
33	148.2	5.0	792	11	BG533818
34	148.2	5.0	952	11	BF793705
35	147	5.0	924	11	BG259918
36	145.8	4.9	709	10	BE382739
37	145.6	4.9	851	11	BG698037
38	140.8	4.8	928	11	BG675913
39	140.2	4.8	528	10	AA450493
40	140.2	4.8	563	10	AA263755
41	140	4.7	844	2	BG779177
42	136.8	4.6	487	10	AW144795
43	136.2	4.6	975	11	BF026745
44	135.6	4.6	688	10	BE729084
45	134.2	4.6	867	11	BG178633

ALIGNMENTS

RESULT 1

AW066305
LOCUS 687009D02.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
DEFINITION 687009D02.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
ACCESSION AW066305
VERSION AW066305.1 GI:6021377
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 653)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 687009. row: D column: 02.
Location/Qualifiers
L. .653
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"
/clone.lib="687 - Early embryo from Delaware"
/tissue.type="embryo"
/dev_stages="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"

AW066305 653 bp mRNA EST 30-MAR-2000
687009D02.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
sequence.
AW066305
AW066305.1 GI:6021377
EST.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 653)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 687009. row: D column: 02.
Location/Qualifiers
L. .653
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"
/clone.lib="687 - Early embryo from Delaware"
/tissue.type="embryo"
/dev_stages="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"

/note="Organ: embryo; Vector: pBluescript SK; Site_1: xhoI
; Site_2: EcoRI; Library was prepared by Stratagene using
the Uni-ZAP XR system (Stratagene BN937328-12). Clones
were picked by a Q-bot after blue/white selection
(ampicillin resistance - use 100 micrograms/microliter).
developed from a pool of equal amounts of RNA from
pollination of the Illinois High Oil Maize Strain Cycle
90. This closed strain has been selected for high oil
concentration for 90 generations and originates from the
1890s era open pollinated variety Burr's White"

BASE COUNT 215 a 98 c 161 g 178 t 1 others
ORIGIN

Query Match 19.7%; Score 579.8; DB 10; Length 653;
Best Local Similarity 96.0%; Pred. No. 7.3e-139;
Matches 616; Conservative 16; Mismatches 6; Indels 4; Gaps 4;

QY 1601 gctactatgtactccagatcattgaacagatgagtggtctgagtgctacgtatttcgta 1660
Db 16 GCTACTATGTACTCCAGATCATTAACAGGATGATGGTCTGAGTCTACGCTATTTCGTA 75

QY 1661 agtgggacggttgagtgagaaattggaggccaaactggaggagatgtcaaaaa 1720
Db 76 AGTGGGACGGTGTGGGAGTGAGAAATTTGGAGGCGCAAAACTGGAGGAGATGTCAAAA 135

QY 1721 ctgaggcaatcaagaattcaaaagattattcttgagaactggaactcatggaag 1780
Db 136 CTGAGGCAATCAAGGAATTCAAAAGATTATTCTTGAGAACACTGGAACTCATGGGAAG 195

QY 1781 ctgggaatgtaaaacaaatttccggaagcgtggagattttaccactgtatgtg 1840
Db 196 CTGGGAATGTAAACCAATTTTCGGAAGCAGCCTGGGAGATTTTACCACCTGATGTTG 255

QY 1841 attatgggttcaagaacacccaaacggaagatatcaatgaaatgaaatgtctcttg 1900
Db 256 ATTATGGTGTAAAGAACGCCAACCGAAAGATATCAGTGAATGAAAGTGTCTCTTG 315

QY 1901 ctctcaattctgaactcatgaagatgcttttcaattgtggagacatatagagctgcta 1960
Db 316 CTCCTCAATTCCTAGACTCATGAAGATGCTTTTCAATGTGGAGACATATAGAGCTGCTA 375

QY 1961 tgatggaattgaaatcaatgtcagaatgctcttgggaagcctaagcmaggraaata 2020
Db 376 TGATGGAATTTGAATTAATATGTCAGAAATGCTCTTGGGAAGCTTAAGCAAGGAAATA 435

QY 2021 ttgaggaagattggaagcattcaactkrgrtcmgratttattggaagcacaccgctnat 2080
Db 436 TTGAGAAAGGATTTGAAGCATTTAATCAGATACAGAAATTTA-TTGAAGGACACCGTGAT 494

QY 2081 caagcactggtctgttgagaagcnaattgttgytggagcmattssytttttcaactct 2140
Db 495 CAAGCCTGTC-TGTTAGAGAAAGCTTAATTTGTTGTCGAGCAATCGCTTTTCACTCT 553

QY 2141 tatectcttctcactctcattatcagggatgagatgttcatattcaaaagcgca 2200
Db 554 TATCCCTTCTATT-ATNCTCATATTATACGGGATGAGGATGATTGATGATCAAA-GCGA 611

QY 2201 aaatgctggaagctctgcaggatattgaaattgcttcaaga 2242
Db 612 AAATGCTTGAAGCTCTGCAGGATATTGAAATTTGCTTCAAGA 653

RESULT 2
AV833309 714 bp mRNA EST 22-JUN-2001
LOCUS AV833309 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
DEFINITION vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA
clone bags5k04, mRNA sequence.
ACCESSION AV833309
VERSION AV833309.1 GI:14525398
KEYWORDS EST.

SOURCE
ORGANISM Hordeum vulgare subsp. vulgare.
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 714)
Sato,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2001)
CONTACT: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kzsato@rib.okayama-u.ac.jp,
URL:http://www.rib.okayama-u.ac.jp/barley/
Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
submission;
database:http://www.shigen.nig.ac.jp/barley/Barley.html.

FEATURES
source
1..714
Location/Qualifiers
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="bags5k04"
/clone_lib="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare shoots germination"
/tissue_type="shoots"
/dev_stage="germination"
BASE COUNT 249 a 121 c 162 g 182 t
ORIGIN

Query Match 16.4%; Score 484.6; DB 10; Length 714;
Best Local Similarity 82.3%; Pred. No. 2.9e-114;
Matches 587; Conservative 19; Mismatches 102; Indels 5; Gaps 5;

QY 1525 cacattcccttragrattgggaaagcatatcacatgcamcttaaacattcttctgacctg 1584
Db 6 CAGGAGGCTGGAAGATGGCAAAAGCATTTCAATACACCCCTAAACAT-TTCTGACATGA 64

QY 1585 cactnaggtgtgacaggtctactatgtatccagatcattgaacagatgtggtctgag 1644
Db 65 CACAAGGTGTTAAACA-GCTACTATATATCTCAGATCATCGAAGAGGATGATGGGAGTGA 123

QY 1645 tgctacgtatttctgaatgggacggttggagtgagaaatggaggcgcaaaactg 1704
Db 124 TCGTATGATTTCGAAAGTGGGGCGAGTTGGCAGTGAAGATTGGTGGAAAGAAACTG 183

QY 1705 gaggagatgtcaaaactgaggaacatcaaggaattcaaaagattattcttctgagaagact 1764
Db 184 GAGGAGATGTCAAAAACTGACGCAATATACATGAATTTAAAGATTATTCTGGAAGAGACT 243

QY 1765 gaaactcatgggaagcttgggaatgttaaaccatttctcgaagcagccttggagatt 1824
Db 244 GAAACCCCTGGGAAGCATGGCAAAAAACAAATTTTCAGAACGACCTCGGGAGATT 303

QY 1825 taccacttgattgttattgttgaagaaacacacacacacacacacacacacacacac 1884
Db 304 TATCCACTTGACATTTGATTAGGAGTTAAGCAAGCAGCAGCAAGCAAGCAAGCATCAGCAA 363

QY 1885 atgaaaagtctctgtctctcactcaatttctagaactcatgaagatgcttttcaattgtgag 1944
Db 364 ATGAAAAGTTCACTTGTCTCTCAGTGTGCAACTCATGATGATGCTTTTCAATGTTGAA 423

QY 1945 acatatagctctcatgtatggaattgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 2004
Db 424 ACATATAGGCTGCTATGATGGAATTTGAAATCAATATGCGCAAGAAATGCCCTTGGGAAA 483

QY 2005 ctgaagcmaggraaattattgagaaagattgaaacttaactkrgrtcmgratttattt 2064
Db 484 TTAAGCAAGGAAAAATATCCAGAAAGGATTTGAAGCATTAATGAGATACAAATCTTA-CT 542

QY 2065 gaaggacacgcgctnatcaagcactgctgtttrgagaagcnaattgttgytsgagcm 2124

```

Db 543 GGATGACACTGGCAATCAAGAACTGGC-TCTTAGAGAGAGCTTGATTGTTGCTGCAAGCA 601
      ||||| | ||||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2125 atssytttttcaactctatccctcttattcatctctcattatatacagggatgagatgatt 2184
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 ATCGTTTCTTCACTCTTATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 661
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2185 tcatattcaaaagcgaaatgcttggaagctctgcagagatattgaaattgcttc 2237
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 TGCAAAATGAAA-GGAAAATGCTTGAAGCTCTTCCAGGATATTGAAATTGCTTC 713
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
BF255013 607 bp mRNA EST 23-FEB-2001
LOCUS HVSMEF0005L18f Hordeum vulgare seedling root EST library HVCDNA0007
DEFINITION (etiolated and unstressed) Hordeum vulgare cDNA clone
HVSMEF0005L18f, mRNA sequence.
ACCESSION BF255013
VERSION BF255013.2 GI:13117445
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 607)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Bequm, D., Frisch, D., Yu
, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and
Wood, T.
Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
On Nov 16, 2000 this sequence version replaced gi:11184130.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7286
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAACTCCTCACTAAAGGG
High quality sequence stop: 551.
Location/Qualifiers
FEATURES
source
1..607
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEF0005L18f"
HVCDA0007 (etiolated and unstressed)
/tissue_type="Seedling root"
/lab_host="TJUC121"
/notes="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"
BASE COUNT 165 a 113 c 163 g 166 t
ORIGIN
Query Match 11.5%; Score 338.2; DB 11; Length 607;
Best Local Similarity 85.6%; Pred. No. 1.9e-76;
Matches 376; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 2509 acgaatttttgggaattcttagcagggcctgaaggaattgcacctctgaggaacctgtt 2568
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACAAAATTTTATGGAATTTCTAGTCAAGGACTAAGGATAGCACCCTCTGAGGCACCCGG 60
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 2569 actggcttatgttcgcaagcctctactttgcagatctagtaagcaagagcgacaa 2628
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AGGGCTATATGTTTGGCAAGGCGCTCTACTTTGCAGATTAGTAAAGCAAGTGCACAG 120
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Qy 2629 tactgttatgttaggaataatcctctgtagttagtcttcttcttcttcttcttcttcttctt 2688
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TATTGTTATGTGGATAGAAAATCGACTGCTGCTTCTTCTGAGCTTCTGCTCTA 180
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2689 ggagacatgtatgaactaaagaagccacgctccatgagcaaacctccaagaggaagcat 2748
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGAGACATGATGACTGAAACCAAGCAAGCCCAATGGACAAACCTCCAAAGAGGAGCAT 240
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2749 tgcaccaagggattagcaaacccgtgcactggagtcagagtttgtgagtgaggagat 2808
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TCACCAAGGGCTTAGGCAAAACACTGTCACATAGAGTCGGAGTTGTTAAATGGAGGAT 300
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2809 gatgtcgtatctccctgcggaagccggtgcccattcatcaattaggagctctcaactcatg 2868
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GATGTCGTCTGCTTGTGGCAAGCCAGTCCAGCATCTATCAGGCATCTCAGCTTCGTG 360
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2869 tacaatgagtcacatctctacaacacatcccaggtgaagatcagttcttctgctgaaggtg 2928
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TACAACGAGTATATAGTGTACAACACAGCTCAGGTGAAGATGACGTTCTTGTGAAGGTC 420
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 2929 cgtttccatcacaagaggt 2947
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AATTCCCTCACAGCGTT 439
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
BF419432 536 bp mRNA EST 24-JUL-2000
LOCUS WWS012.A12R000101 ITEC WMS Wheat Scutellum Library Triticum
DEFINITION aestivum cDNA clone WWS012.A12, mRNA sequence.
ACCESSION BF419432
VERSION BF419432.1 GI:9417278
KEYWORDS bread wheat.
SOURCE Triticum aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 536)
Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier
, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
Pecchioni, N., Quaslet, C., Schuch, W., Selvaraj, G., Shariflou, M.,
Sorrenti, M., Warburton, C. and Wenzel, G.
International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Schuch W
Zeneca Wheat Improvement Centre, Norwich Research Park
Colney Lane, Norwich NR4 7UH UNITED KINGDOM
Tel: 44 1603 250 2600
Fax: 44 1603 250 699
Email: wolfgang.schuch@aguk.zeneca.com
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
FEATURES
source
1..536
/organism="Triticum aestivum"
/cultivar="Novosibirskaya 67"
/db_xref="taxon:4565"
/clone="WWS012.A12"
/clone_lib="ITEC WMS Wheat Scutellum Library"
/tissue_type="scutellum callus"
/notes="M13 Reverse sequencing primer used for 5' end of
clone."
BASE COUNT 143 a 106 c 144 g 139 t 4 others
ORIGIN
Query Match 10.6%; Score 312; DB 10; Length 536;
Best Local Similarity 85.1%; Pred. No. 1.1e-69;

```

25 microm ABA (abscisic acid) in 5 mM Mes buffer, pH 5.7, for 12 hr at 22 C. The tissue, total RNA, and poly(A) RNA were prepared by Steven Verhey in M.K. Walker-Simmons's lab (USDA-ARS, Washington State Univ., Pullman, Washington 99164-6420. A cDNA library was made by Clontech using a combination of random and oligo dt primers. Library was plated and archived by Russell Johnson (Colby College, ME/Walker-Simmons' lab). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT	122 a	93 c	84 q	128 t

[illegible]

57 ATCTCCATTAACAAGATGCTATTATGGCATGGTTC AAGGCTG 17

RESULT	6	EST	22-JUN-2001
AV834168/c	AV834168	661 bp	EST
LOCUS	AV834168	661 bp	EST
DEFINITION	K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA clone rbags5k04, mRNA sequence.		

DEFINITION
AV034108 N. satou unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA clone rbags5k04, mRNA sequence.

EST.	REINWORDS	SOURCE	ORGANISM
	Hordeum	vulgare subsp. vulgare.	
	Hordeum	vulgare subsp. vulgare	
	Eukaryota;	Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta;	Magnoliophyta;	Liliopsida; Poales; Poaceae; Pooidae
	; Triticaceae;	Hordeum.	

REFERENCE
1 (bases 1 to 661)
; ATCC; NORDELL.

AUTHORS Sato, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2001)
COMMENT Contact: Kazubiro Sato

Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan

Email: kazzsato@rib.okayama-u.ac.jp,
URL: <http://www.rib.okayama-u.ac.jp/barley/>
Sato, K., Saisho, D., Takeda, K., Shini, T. and Kohara, Y. Direct
submission:

Fri Mar 8 08:31:54 2002

```

Db 538 TTGAGAGACTGGCACTCATGGAGCATGGCAACAAAAGAAAAATTTTCAGAACCAAC 597
Qy 1814 ctggagatttaccacacttgatttgattatgtgttgaagaagaccacaaacggaag 1873
Db 598 CTCGCAGATTTTCTCTCGATATTGATTTATGGGGTGGACAAAAAACCCCTCTAAG 657
Qy 1874 a 1874
Db 658 A 658

RESULT 10
BG042229 502 bp mRNA EST 31-JUL-2001
LOCUS BG042229.1 Gm-cl055 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl055-2064 5' similar to TR:Q92P54 Q92P54 POLY(ADP-RIBOSE)
POLYMERASE ;, mRNA sequence.
ACCESSION BG042229
VERSION BG042229.1 GI:12488706
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 502)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 429 Std Error: 0.00
High quality sequence stop: 419.
FEATURES
Location/Qualifiers
1..502
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl055-2064"
/clone_lib="Gm-cl055"
/tissue_type="Mature seed pods, greenhouse grown"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site-1: EcoRI; Site-2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature seed pods of greenhouse grown plants prior to
senescence for the cultivar Kpl. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments
were transformed into DH10B host cells (GibcoBRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker."
150 a 80 c 119 g 153 t

BASE COUNT
ORIGIN

```

```

Query Match 6.4%; Score 189; DB 11; Length 502;
Best Local Similarity 73.0%; Pred. No. 6.3e-38;
Matches 243; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
Qy 2615 gcaagagcgcaactactgttattgttgtagtaataatccttaggtttgtatgctcttt 2674
Db 2 GCAAAAGTGTCTAGTATGCTTCACTGATGAAGAAAAATCCTGTGCTCTAATGCTTTTGA 61
Qy 2675 ctgaggttcttttagggagacatgtatgaactaaaagaagccacgtccatggacaacctc 2734
Db 62 GTGAAGTTGCGCTTGGAAATGTCTATGAGCTCAAGAAAGCTAAGTATATGGTAACCTC 121
Qy 2735 caagagggagcatttcgacccaaggattagcacaacccgtgcactggagtcagagtttg 2794
Db 122 CCGAAGGAAGACACCTCTACTAAAGGACTGGCAAGAAATGCCACAGGAATCGGAATATG 181
Qy 2795 tgaagtggaggatgatgtctgttagtccctcgggcaagccgggtgccatcatcaattagga 2854
Db 182 TAAAGTGGAGGGCAATGTCTACTGTTCCTTGTGGCAACACAGTGCCTCAAAATGTCAAGA 241
Qy 2855 gctctgaactcatgtacaatgaatgactacatcgtctacacacatcccccaggtgaagatgcag 2914
Db 242 GTTCTGAGCTCATGTCAATGAGTATATGTTTATACTGCTCAAGTTAAGATGCAAT 301
Qy 2915 tctgctgaaggtgcgttccatccatcaagaaggt 2947
Db 302 TCITATTGAAGGTGAGGTTTCATCACAAGAGAT 334

RESULT 11
LOCUS AQ577141 759 bp DNA GSS 02-JUN-1999
DEFINITION nbxb0090B07r CUGI Rice BAC Library Oryza sativa genomic clone
nbxb0090B07r, DNA sequence.
ACCESSION AQ577141
VERSION AQ577141.1 GI:4977626
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 759)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACACGATATGACCATG
Class: BAC ends
High quality sequence stop: 404.
Location/Qualifiers
1..759
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbxb0090B07r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of

```

carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening." 148 c 174 g 207 t 1 others

BASE COUNT	229 a	148 c	174 g	207 t	1 others
ORIGIN	One whole library for colony screening. "				
Query Match	6.3%; Score 186.2; DB 13; Length 759;				
Best Local Similarity	74.9%; Pred. NO. 3.7e-37;				
Matches 233; Conservative	0; Mismatches 78; Indels 0; Gaps 0;				
QY	570	taagggttccaagcgcaagaaagtgaatgatattgtagctacaaatcgccaggtt	629		
Db	236	TAAGGTTTCCAAGCGCAAGATAATCAGATGATATACAGACTCCAAAGCTCCCAAGT	295		
QY	630	agatgaagtagcacatctgaaggtacagtggaaacaaagggcaactttagaccacgtgg	689		
Db	296	AATCAGGAGTATATTCAGAAGGCACAGCAAGATAAAGGGAAGCTGTTGATCACATGA	355		
QY	690	ttccaaactactagttccagctgataccaaactaaagcttaaggagcaaaagcacacttg	749		
Db	356	TTCCAATGCTAAATTCCTTCATCTTCCAGAAAGCTTTAAAGAGCAGAGTGCACACTTTG	415		
QY	750	gaagttaaagatgtagcttaagactcatgtatcgctgctgaattaaaggatattgctga	809		
Db	416	GAATTTGAAGGATGAACCTTAAAGAAGCATGTGCTCAACTGCTGAACTAAGGAACATGCTCGA	475		
QY	810	ggctaaatggcgagatcacatcaggacacgaagaagggcacctattgtagctgtgcggatgg	869		
Db	476	AGCAATGGCGAGGATACATCTGGGCGCAGGGCATCTATTAGACCGTTGGTAATATTG	535		
QY	870	aatgctatttg	880		
Db	536	CATGCTATTTG	546		

RESULT	12
LOCUS	AA401836
DEFINITION	zv66b12.rl Soares.Total_fetus_NB2HF8_9w Homo sapiens cDNA clone IMAGE:758591 5' similar to gb:J03473 NAD(+) ADP-RIBOSYLTRANSFERASE (HUMAN); mRNA sequence.
ACCESSION	AA401836
VERSION	AA401836.1
KEYWORDS	GI:2055855
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 669)
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubouque,T., Geisels,G., Jost,S., Kucaba,T., Lacy,M.M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Willson,R.
TITLE	WashU-Merck EST Project 1997
JOURNAL	Unpublished (1997)
COMMENT	Contact: Willson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: estewartson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ~28ml3 rev2 ET from Amersham
 High quality sequence stop: 493.

	FEATURES	SOURCE
1.	100% pure	100%
2.	100% pure	100%
3.	100% pure	100%
4.	100% pure	100%
5.	100% pure	100%
6.	100% pure	100%
7.	100% pure	100%
8.	100% pure	100%
9.	100% pure	100%
10.	100% pure	100%
11.	100% pure	100%
12.	100% pure	100%
13.	100% pure	100%
14.	100% pure	100%
15.	100% pure	100%
16.	100% pure	100%
17.	100% pure	100%
18.	100% pure	100%
19.	100% pure	100%
20.	100% pure	100%
21.	100% pure	100%
22.	100% pure	100%
23.	100% pure	100%
24.	100% pure	100%
25.	100% pure	100%
26.	100% pure	100%
27.	100% pure	100%
28.	100% pure	100%
29.	100% pure	100%
30.	100% pure	100%
31.	100% pure	100%
32.	100% pure	100%
33.	100% pure	100%
34.	100% pure	100%
35.	100% pure	100%
36.	100% pure	100%
37.	100% pure	100%
38.	100% pure	100%
39.	100% pure	100%
40.	100% pure	100%
41.	100% pure	100%
42.	100% pure	100%
43.	100% pure	100%
44.	100% pure	100%
45.	100% pure	100%
46.	100% pure	100%
47.	100% pure	100%
48.	100% pure	100%
49.	100% pure	100%
50.	100% pure	100%
51.	100% pure	100%
52.	100% pure	100%
53.	100% pure	100%
54.	100% pure	100%
55.	100% pure	100%
56.	100% pure	100%
57.	100% pure	100%
58.	100% pure	100%
59.	100% pure	100%
60.	100% pure	100%
61.	100% pure	100%
62.	100% pure	100%
63.	100% pure	100%
64.	100% pure	100%
65.	100% pure	100%
66.	100% pure	100%
67.	100% pure	100%
68.	100% pure	100%
69.	100% pure	100%
70.	100% pure	100%
71.	100% pure	100%
72.	100% pure	100%
73.	100% pure	100%
74.	100% pure	100%
75.	100% pure	100%
76.	100% pure	100%
77.	100% pure	100%
78.	100% pure	100%
79.	100% pure	100%
80.	100% pure	100%
81.	100% pure	100%
82.	100% pure	100%
83.	100% pure	100%
84.	100% pure	100%
85.	100% pure	100%
86.	100% pure	100%
87.	100% pure	100%
88.	100% pure	100%
89.	100% pure	100%
90.	100% pure	100%
91.	100% pure	100%
92.	100% pure	100%
93.	100% pure	100%
94.	100% pure	100%
95.	100% pure	100%
96.	100% pure	100%
97.	100% pure	100%
98.	100% pure	100%
99.	100% pure	100%
100.	100% pure	100%

```
1. .669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:759591"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10R"
```

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5',
 TGTTTCAATCTGATGGAGCGCGCTTAATTTTTTTTTTTT 3',
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo. "
 191 a 152 c 157 q

BASE COUNT
ORIGIN

Query Match 6.1%; Score 181; DB 10; Length 669;
Best Local Similarity 56.5%; Pred. No. 7.9e-36;
Matches 378; Conservative 0; Mismatches 285; Indels 6; Gaps 2;

QY	2276	ttgatataaatatgaaacttcacttgacatcaccccgctgcactcagatagtgaag	2335
Db	3	TCGATCTCAACTATGAGAAGCTCAAAACTGCATTAAGTGGTTGACAGAGATTCCTGAAG	62
QY	2336	attacaagttaattgagcagtgatcctccaacacacatgctccctactcaaaaggaactggt	2395
Db	63	AAGCCGAGATCAFCAGGAAGTATGTTAAGAACACTTCATGCACCAACACACAAATCGGTATG	132
QY	2396	cgcgtgaactggaggaagttttttcacttgatcagatgagagagaaacttaataagtactcaa	2455
Db	123	ACTTGGAAAGTCATCGATATCTTTTAAGATGAGACGCTGAAGGCGAATGATCGCTTACAAAGC	182
QY	2456	gatataaaataatctgcataaacaagatgctattatggcacggtgtccaaaggttgacgaatt	2515
Db	183	CCTTT---AAGCAGAGCTTCATAACCCGAAGATTGCTGTGSCACGGGTCCAGCACCCCACT	239
QY	2516	tgtgtgggaattcttagtcaaaaggctaaagaattgcactcctcgaggaacctgttactggct	2575
Db	240	TTGCTGGGATCTGTCCACAGGCTTTCGGATAGCCCCGCTGAAGGCCCGGTGACAGGCT	299
QY	2576	atatgtcggcaaaagccctactttgcagatctagtaagaagaagcgacacatactgtt	2635
Db	300	ACATGTTGGTAAAGGGATCTATTTCGTGACATGGTCTCAAGAGTGCACCACTACTGCC	359
QY	2636	atgtggaaggaaataactcctgtaggtttgatgctctcttcgaggtgtcttagagaca	2695
Db	360	ATACGCTCTCAGGGAGACCAATAGGCTTAATCCTGTGGGAGAGTGTCCCTTGGAAACA	419
QY	2696	tgtatgaactaagaagaagcaagtcocatgagacaacctccaagagggaagcatcgacca	2755
Db	420	TGTATGAACCTGAAGCAGCGTTTCATATCAGCAAGTATACCCAAAGGCAAGACAGCTGTC	479
QY	2756	agggattaggcaaaaccgctgccactgagtcagagtttgtgaagtggagggatgatgtcg	2815
Db	480	AAGGTTTGGGCAAAACTACCCCTGATCCTTCAG--CTAACATTAGTCTGGATGGTGTAG	536
QY	2816	tagttccctgcgcgaacgggtccatcatcaataggagcctgaactcatgtacaatg	2875
Db	537	ACGTTCCTCTTGGGACCGGGATTTCACTCGGTGTAATGACACCTCTCTTACATATATAACG	596
QY	2876	agtcacatgctacaacaacatcccaagtggaagatgcagttcttgtcgtaaggtgcgtttcc	2935
Db	597	AGTACATTGCTATGATATTGCTCAGGTAATCTGAAGTATCTGCTGAACTGAACTTCA	656

296 AGCGTGAAGCGGAATGCCAGCGTTTACAAGCCCTTT---AAGCAGCTTCATACCGGAGAT 352
 2486 tattatggcaacggttcaaggttgacgaatttgggaattcttagtcaagggtcaagaa 2545
 353 TGCCTGTGGACGGGTCCAGGTCACACCAACTTTGCTGGGATCTGTGCCAGGGTCTTCGGA 412
 2546 ttgcacctctgagcgaccctgttactgcttatgttctggcaaaagcctctactttcag 2605
 413 TAGCCCCGCTGAAGCGCGGTGACAGCTACATGTTGGTAAAGGGATCTATTTTCGCTG 472
 2606 atctagtaagcaagcgacaaatactcttatgttgataggaatactctgtagtttga 2665
 473 ACATGGTCTCCAAGAGTGCCACTACTGCCATCGTCTCAGGGAGACCCAATAGGCTTAA 532
 2666 tgcctcttctgaggttgcttttaggagacatgatatgaactaaagaagccacgctccatgg 2725
 533 TCCTCTGGGAGAGTTGCCCTTGGAAACATGTTATGAACTAAGCAGCGTTCACATATCA 592
 2726 acaaacctccaagaggaagcattcgacagggattagcgaacacggtgccactggagt 2795
 593 GCAAGTTATCCCAAGGGCAAGCAGTGTCMAAGGTTGGGCAAACTACCCCTGTATCCTT 652
 2786 cagagtttgtaagtgagggtatgctgtagttccctgcggcaagcgggtgcatcat 2845
 653 CAG---CTAACATTTAGTCTGAGTGGGTAGACGATCTCTTTGGGACCGGATTTCACTGG 709
 2846 caattaggagctctgaaactaatgataatgagtacatcgctacacacatccacggtga 2905
 710 GTGTGAATCACACT-CTCTACTATATACGAGTACATTTGCTATCATATCTGTCAAGTAA 768
 2906 agatgcagttcttgcgaagtcggtttccatccacaaagggt 2947
 769 ATCTGAAGTATCTGTGTGAACACTGAAATTCAAATTTTAAGAGCT 810

RESULT 14
BI093436 916 bp mRNA EST 20-JUN-2001
602859629f1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5000911 5',
DEFINITION
mRNA sequence.
ACCESSION
BI093436
VERSION
BI093436.1 GI:14511766
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 916)
REFERENCE
NTH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
robe-s@mail.nih.gov

CDNA Library attached to the
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11033 row: 1 column: 08
High quality sequence stop: 837.
Location/Qualifiers
1. .916
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5000911"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: Not
Site_2: Sall; Cloned unidirectionally. Primer: Oligo (

FEATURES
source[illegible]

Average insert size 1.5 kb. Library prepared by Life Technologies."	
248 a	209 c 230 g 229 t
BASE COUNT	
ORIGIN	

Query Match	5.9%	Score 174.6	DB 11	Length 916
Best Local Similarity	56.4%	Pred. No. 3.8e-34		
Matches 368	Conservative	0	Mismatches 279	Indels 6
Gaps				
2292	gaaacttcaactgtgaacatcaccccgctggctccacgtagtgaagattacaaggttaattga	2351		
4	GAAGCTCAAAACTGACATTAAGGTGGTTGACAGAGATTCTGAAGAAGCCGAGATCATCAG	63		
2352	gcagttatctctcaacacacatctctcctactcaacgaagactggctcgtcgtggaactggagga	2411		
64	GAAATATGTTAAAGAACACTCATGCAACCCACACACAATCGGTATGACTTTGGAAGTCATCGA	123		
2412	agtttttctacttgatcgagatggagaacttaataagttactcaagatatataaaataatct	2471		
124	TATCTTTAGATAGACGGTGAAGGCCGAATGCCAGCGTTTACAAGCCCTTTAGCAG	---CT 180		
2472	gcatacaagatgctattatggcacggttcctcaaggttgacgaattttg99aaattcttag	2531		
181	TCATAACCGAAGATTGCTGTGSCACGGGTCAGGACCACCAACTTTCGTGGGATCTCTGTC	240		
2532	tcaagggtcaagaattgcactcctcctgagccactgttactggcttatgttcggcgaagg	2591		
241	CCAGGGTCTTCGGATAGCCCGCCCTGGAAGCCCGGTGACAGGCTACATGTTTGGTAAAGG	300		
2592	cctctactttgcagatctagTaaGcaagagcgcaatactgttatgtg9ataggaataa	2651		
301	GATCTATTTCGTGCATAGTGGTCTCCAGAGTGCCCACTACGCTACGCTCAGGGAGA	360		
2652	tctgtaggtttgatgcttcttctggagttgcttttaggagacatgtatgaactaaagaa	2711		
361	CCCAATAGGCTTAATCTGTGGGAGAAGTGTGCCCTTGGAAACATGTATGAATGAAGCA	420		
2712	agccacgtccatggacaaacctccaagaggggaagcattcgaccaaggaattaggcaaac	2771		
421	CGCTTCATATCAGCAAGTTATCCCAAGGGCAAGCACAGCTGTCAAGGTTTGGGCCAAAC	480		
2772	cgtgccactggatcagagtttgtgaagtggagggaigatgctgtagttccctcgcgcaa	2831		
481	TACCCCTGATCCTCAG---CTAACATTTAGCTGTGATGGGTGTAGACGTTCTCTTGGGAC	537		
2832	gcgggtgccaatacaattaggagctctgaaactcatgtacaaatgagtcacatcgctcaaa	2891		
538	CGGATTTTCATCTGGTGTGAATGACACCTCTCTACTATATTAACGAGTACATTGTCTATGA	597		
2892	cacatcccagggtgaagatcgagttcttctgaaggtcggtttccatcacaa	2944		
598	TATTTGTCAGGTAATCTGAAGTATCTCGTGAACGTGAATCTCAATTTTAAAG	65		

RESULT	15
BG031594	
LOCUS	990 bp mRNA
DEFINITION	G03299739F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4394288 5', EST
ACCESSION	BG031594
VERSION	BG031594
KEYWORDS	BG031594.1 GI:12422028 EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 990)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgabbs-re@mail.nih.gov
COMMENT	

Tissue Procurement: DCTD/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L7A1M10089 row: m column: 09
High quality sequence stop: 678.

FEATURES
SOUND

```

1. 3250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4394288"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
274 a 218 c 251 g 247 t

```

BASE COUNT
ORIGIN

Query Match	5.8%;	Score 170.8;	DB 11;	Length 990;
Best Local Similarity	55.7%;			
Matched	280	Pred. No. 3.7e-33;		

	Conservative	0	Mismatches	302	Indels	7	Gaps	3
QY	2247	tgcttogatgcagcagtgatgaatctcttgatgataataatatagaacttactgtga	2306					
Db	8	TGGGTCTGATGATAGCAGCAAGGATCCCATCATGTCAACTATGAAAGGTCFAAACTGA	67					
QY	2307	catcacccgcctggctccagcagtagtgaagattacaagttaattgagcagtatctctcaa	2366					
Db	68	CATTAAAGGTGGTTGACAGAGATTCGAAAGACCGAGATCATCAGGAAGTATGTTAAGAA	127					
QY	2367	cacacatgctctactcaacaggactggctgcctggaactggaggaagcttttttcaactga	2426					
Db	128	CACTCATGCAACACACACAATGGGTATGACTTGGAACTCATCGATATCTTTAAGATAGA	187					
QY	2427	tcgagatggagaacttaataagttactcaagataataaaataatctgataacaagatgct	2486					
Db	188	GGGTGAAGGCGAATGCCAGCCTTACAGCCCTTT--AAGCAGCTTCATAACCCGAAGATT	244					
QY	2487	attatggcacaggttcaagggtgaacgaattttgtgggaattcttagtcaaggggctcaagaat	2546					
Db	245	GCTGTGGCACGGGTCCAGGACCCACCAACTTTGTGTGGGATCCTGTGCCAGGGCTTCGGAT	304					
QY	2547	tgcacctctgagcgacctgttactggtctatatgttcggcaaggcctctactttgcaga	2606					
Db	305	AGCCCGCGCTGAAGCGCCGTGACAGGCTACATGTTTGGTAAGGGGATCTATTTCGCTGA	364					
QY	2607	tctagtgaacaagagcacaactctgttatgtggaatgaataatcctgtaggtttgaat	2666					
Db	365	CATGGTCTCCAAGAGTGGCCAACTACTTGCATACGTCCTCAGGGAGACCCAATAGGCTTAAT	424					
QY	2667	gctctcttgaggttgcttttaggagacatgtatgaactaaagaagcacaagtcgaatgga	2726					
Db	425	CCTTGTGGGAGAAAGTTGCCCTTGGAAACATGTATGACTGAAGCACGCTTCACATATACG	484					
QY	2727	caaacctccaagaggaagcattccaccaagggattaggcaaaaccgtgccactggagtc	2786					
Db	485	CRAAGTTACCCAGGCGACGACAGTGTCAAAGGTTTGGCAAAAACACTACCCTGATCCCTTC	544					
QY	2787	agagtttgaagtggagggatgatgtcgttagttccctcgcggaagccggtgcacatc	2846					
Db	545	AG---CTAACATTAGCTCGGATGGTGTAGAGTTCCTCTTGGGACCGGGATTTCATCTGG	601					
QY	2847	aattaggagctctgaactcatgtacaatgagtacatcgtctacaacacatcccaggtgaa	2906					
Db	602	TGTGAATGACACCTCTTACTATATACGAGGTACATTGTCTATGATATTGCTCA-CTACA	660					

~~us-09-236-995d-1.rst~~

Fri Mar 8 08:31:54 2002

Qy 2907 gatcagattcttgctgaagggtggtttccatcaaga 2944
 ||| ||||| | ||| |||
Db 661 TCTGAGTATCTGCTGAACACTTCAATTCAATTGAAGA 698

Search completed: March 7, 2002, 15:24:26
Job time: 3336 sec

